ACTA LINGUISTICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS XXXVI FASCICULI 1–4

AKADÉMIAI KIADÓ, BUDAPEST

1986

ACTA LINGUISTICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Acta Linguistica publishes papers on the subjects of Finno-Ugric, Slavonic, Germanic, Oriental and Romance linguistics as well as general linguistics in English, German, French, Russian and Italian.

Acta Linguistica is published in yearly volumes of four issues by

AKADÉMIAI KIADÓ

Publishing House of the Hungarian Academy of Sciences
H-1117 Budapest, Prielle Kornélia u. 19–35

Manuscripts should be addressed to:

ACTA LINGUISTICA
H-1014 Budapest, Szentháromság u. 2

Correspondence with the editors and publishers should be sent to the same address.

Orders may be placed with

AKADÉMIAI KIADÓ
H-1519 Budapest, P.O. Box 245

Acta Linguistica is abstracted/indexed in Current Contents–Arts and Humanities, Arts and Humanities Citation Index

NOTICE TO AUTHORS

To facilitate editorial work and to enhance the uniformity of presentation, authors are requested to prepare their contributions in accordance with the conventions summarized below.

Form of manuscript

The language of publication may be English, French, German or Russian. Please provide typewritten double-spaced manuscripts with wide margins (50 characters per line, 25 lines per page). Type on one side of the paper only and number all the pages consecutively. The following information should be provided on the first page: the title of the article (or, in case of reviews, the title of the book under discussion, the name of

ACTA LINGUISTICA

ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS XXXVI FASCICULI 1–4

AKADÉMIAI KIADÓ, BUDAPEST

1986

ACTA LINGUISTICA

TOMUS XXXVI

INDEX

Beard, R.: Neurological evidence for lexeme-morpheme base morphology 3
Botha, R. P.: The role of conceptualization rules in the interpretation of morphologically complex words 25
Brekle, H. E.: The production and interpretation of *ad hoc* nominal compounds in German: A realistic approach 39
Dressler, W. U. and *Merlini Barbaresi, Lavinia:* How to fix interfixes? On the structure and pragmatics of Italian (and Spanish, Russian, Polish) antesuffixal interfixes and of English "intermorphic elements" 53
Groot, C. de: Predicate formation in functional grammar 69
Kastovsky, D.: Problems in the morphological analysis of complex lexical items 93
Kenesei, I.: On the role of the agreement morpheme in Hungarian 109
Marle, J. van: Lexical mechanisms versus morphological structure..... 121
Meijs, W.: Over the borderline: Words or phrases? 135
Pordány, L.: Irreversible binomials in Hungarian 151
Pounder, Amanda: The semantic organization of word-formation paradigms and diachrony 181
Rainer, F.: Recursiveness in word-formation, with special regard to Spanish . 197
Toman, J.: Transparent head, inheritance and the normal form 211
Ulrich, Miorita: Thetisch vs. kategorisch und Informationsstruktur 221
Varela, Soledad: The organization of the lexical component: Noun-compounds in Spanish 235
Vogel, Irene: Phonological evidence for level ordering in Italian word formation 245
Zwanenburg, W.: Derivation constructions and argument structure 261

CRITICA

L. Bouves: The Phonetic Basis of Perceptual Ratings of Running Speech (*Mária Gósy*) 273
Dictionary of American Regional English, Volume I. Chief Editor: F. G. Cassidy (*M. Kontra*) 275

INDEX

Beard, R.: Neurological evidence for lexeme-morpheme base morphology 3
Botha, R. P.: The role of conceptualization rules in the interpretation of morphologically complex words 25
Brekle, H. E.: The production and interpretation of *ad hoc* nominal compounds in German: A realistic approach 39
Dressler, W. U. and *Merlini Barbaresi, Lavinia:* How to fix interfixes? On the structure and pragmatics of Italian (and Spanish, Russian, Polish) antesuffixal interfixes and of English "intermorphic elements" 53
Groot, C. de: Predicate formation in functional grammar 69
Kastovsky, D.: Problems in the morphological analysis of complex lexical items 93
Kenesei, I.: On the role of the agreement morpheme in Hungarian 109
Marle, J. van: Lexical mechanisms versus morphological structure 121
Meijs, W.: Over the borderline: Words or phrases? 135
Pordány, L.: Irreversible binomials in Hungarian 151
Pounder, Amanda: The semantic organization of word-formation paradigms and diachrony 181
Rainer, F.: Recursiveness in word-formation, with special regard to Spanish . 197
Toman, J.: Transparent head, inheritance and the normal form 211
Ulrich, Miorita: Thetisch vs. kategorisch und Informationsstruktur 221
Varela, Soledad: The organization of the lexical component: Noun-compounds in Spanish 235
Vogel, Irene: Phonological evidence for level ordering in Italian word formation 245
Zwanenburg, W.: Derivation constructions and argument structure 261

CRITICA

L. Bouves: The Phonetic Basis of Perceptual Ratings of Running Speech (*Mária Gósy*) 273
Dictionary of American Regional English, Volume I. Chief Editor: F. G. Cassidy (*M. Kontra*) 275

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 3–23 (1986)

NEUROLOGICAL EVIDENCE FOR LEXEME-MORPHEME BASE MORPHOLOGY

R. BEARD

(Bucknell University, USA)

1. The Type Transparency Hypothesis

Since its inception, the generative school of linguistic theory has struggled to establish an attitude toward the relation of its competence model of grammar to a psychological model of performance (parser). Initially it was hoped that the relation would be "transparent", that the items and operations of grammar would be isomorphic with those of the parser. However, early attempts at confirming the psychological reality of transformations came to naught and Chomsky retracted his position to one which allowed the opaque, homomorphic relationship developed by Fodor, Bever and Garrett 1974.

More recently Bresnan (1978) and Berwick and Weinberg (1983, 1984) have developed an argument for using the performance model as a tool for the evaluation of the adequacy of theories of grammar, claiming that the more transparent the relation of grammar to parser, the more adequate the model. Berwick and Weinberg call this the "Type Transparency Hypothesis". Bresnan feels that in order for type transparency to work for grammar we must abandon transformation rules in the absence of evidence that they operate directly in speech and in view of the linear operation of parsers. However, Berwick and Weinberg demonstrate that Bresnan's Lexical Function Grammar, which accounts for transformational relations in the lexicon, is superior by this test only assuming a cognitive mechanism capable of carrying out a single operation at a time. If we assume a more reasonable model of human mental capacities, a mechanism which can carry out several operations simultaneously, the various levels of ST, EST, REST and even GB can all be accomodated in a transparent parser. They propose such a model, following lines originally laid out by Marcus (1980).

In this paper I will demonstrate that given these latest assumptions about the nature of mental processing, the morphological model which receives the greatest support from the neurological evidence is one which separates morphology from the lexical and syntactic operations it marks (the **Separation Hypothesis** (SH)) and distinguishes lexemes and (grammatical) morphemes as discrete natural classes situated in autonomous components of grammar (the **Lexeme-Morpheme Base Hypothesis** (L-MBH)). Section 2 outlines such a model.

Acta Linguistica Academiae Scientiarum Hungaricae 36, 1986
Akadémiai Kiadó, Budapest

Section 3 reviews the "classical theory" of aphasia, a syndrome of speech failure associated with distress of the speech regions of the left cerebral cortex. Section 4 then demonstrates type transparency between this model and the parser with evidence from research in aphasia. It then goes on to show how one series of recent attacks on the classical theory of aphasia suggests extensions of that theory which in fact strengthen the aphasiological support for SH and L/MBH.

2. The Grammatical Model

2.1. The Separation Hypothesis. Recent interest in linguistic (Chomsky 1981) and cognitive modularity (Fodor 1983) has influenced a growing sympathy for separating morphology from lexical and syntactic derivation in the literature (Anderson 1982, Beard 1976, 1981, Klavans 1985, Pesetsky 1985, Pounder 1986, Sadock 1985, Szymanek 1985). In my version of the Separation Hypothesis, morphemes are purely phonological operations independent of any and all lexical or syntactic operations on grammatical functions of the stem.

The best argument for this hypothesis is the fact that the conditions on lexical and syntactic derivation are independent of those on affixation. Conditions on morphology are related to derivations and the grammatical features they manipulate only in that such features, added previously to a lexeme, form some part of the conditions on morphological marking. The English subjective (agentive) nominalization seems to operate on verbs with few real restrictions. A small, lexically marked group of mostly intransitive verbs, e.g. the copulas *be, seem, become* and a few others, *stay, fall, weigh*$_{[-Tr]}$: **be-er, *seem-er, *becom-er, *stay-er, *fall-er* (*?fall-ee*), **weigh*$_{[-Tr]}$*-er* seem semantically prohibited from subjectivization. However, neither intransitive

(1) *whin-er* *revel-er* *runn-er* *grovel-er* *snivel-er*

nor nonagentive verbs

(2) *sleep-er* *dream-er* *bleed-er* *experienc-er* *remain(d)-er*

are strictly excluded. Acceptibility depends on the availability of a generic referent, i.e. a class of conceivable referents which the derivate might type name, but the output potential seems to depend on no more than the capacity of the underlying verb to support a subject (**rain-er, *snow-er*). In any event, marginally attested forms and related compounds like *?com-er* but *late-comer*, *?go-er* but *church-go-er* demonstrate that the restrictions on such usage are unrelated to the ability to combine with *-er*.

Marchand (1969) has further undermined the argument that the lexical derivation rule here generates agentive rather than subjective nominalizations with evidence of wide-spread inanimate subjective derivations in English.

(3) *trail-er* N *trails* (the feature film, an automobile)
sparkl-er N *sparkles*
sink-er N *sinks* (weight, baseball pitch)
thrill-er N *thrills* (the experiencer)

Assuming that these derivations cannot be instrumental nominalizations, they must be the grammatical outputs of a subjective not agentive nominalization.

Constraints on affixation are more elaborate and different from those on derivation. Initially, it might seem that affixation parallels derivation in that all subjective derivations tend to receive *-er* regardless of transitivity. Intransitive Latinate verb stems, however, consistently receive the affix of the agentive adjective whatever it might be.[1]

(4) Latinate (Stative) Intransitive Subject Nominalizations

	Object Case		Object Case
Derivate	Marker	Derivate	Marker
particip-ant	(in)	*(e/im)migr-ant*	(from/to)
adher-ent	(to)	*descend-ent*	(from)
registr-ant	(for)	*degener-ate*	(from/to)
matricul-ant	(at)	*devi-ate*	(from)
resid-ent	(at)	*operat-ive*	(in)

Subjective nominalizations based on intransitive, especially stative verbs, presently tend to receive *-ee*, the same suffix used to mark animate (direct and indirect) objective nominalizations.[2] This suggests that affixation alone is determined by agentive and patientive features, without reference to the subjective/objective derivational distinction.

(5)	*stand-ee*	(Ø)	*wait-ee*	(for)
	escap-ee	(from)	*retir-ee*	(from)
	embark-ee	(on)	*resign-ee*	(from)
	adapt-ee	(to)	*confer-ee*	(with)

[1] The number of exceptions is remarkably small, mostly in the direction of the intransitive marker marking transitive verbs: *student, assistant* (but note *assistant to the manager*), *disscussant*. The few intransitive Latinate verbs not receiving the now productive *-ee* but receiving *-er* either end on *-nt* already (*dissenter*) or have become germanicized (*reveller*, cf. **revelant*).

[2] See Beard (1986) for a discussion of the autonomy of conditions on objective derivations and the suffixes, *-ing*, *-ee*, etc. which mark them.

Finally, if the underlying verb has an associated particle, no affix is added to the underlying stem.[3]

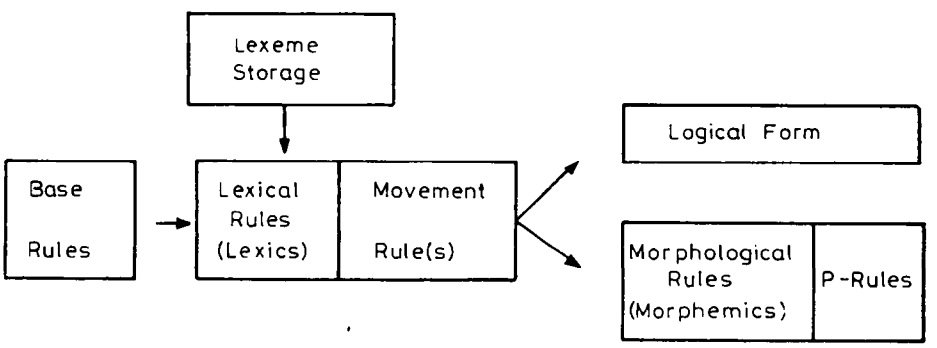
(6) *a cut up* *a drop out* *a look out*
a stand in *a walk in* *a stand by*
a show off *a run away* *a hold out*

Notice that even if this rule is unproductive, the argument goes through; (6) demonstrates that omission would be the preferred marking regardless of the status of the derivation rule. This marking is contingent on the presence of the particle, not the derivation, for all count nominalizations of verb-particle couples are omissively marked, e.g. these both locative and prefective nominalizations *a get-away*, *a turn-around* and *a drop-off*, *walk-in*, which are also subjectives.

The contemporary morphology for the subjective nominalization requires (i) the corresponding adjective affix for lexically marked intransitive Latinate verbs.[4] Otherwise, (ii) the subjective nominalization of intransitive (stative) verbs is productively marked by the suffix *-ee*, which as Carrier-Duncan (1985, 32—3) points out, is an animate "nonagentive" marker. Germanic verb-particle stems (iii) are marked with omissive morphology and (iv) elsewhere the suffix *-er* marks the subjective nominalization in English. None of these conditions correspond to the conditions on the subjective or, for that matter, any putative agentive nominalization. All the lexical derivates of (1—6) may be generated by a single subjective nominalization rule with virtually no constraints; however, the constraints on affixation are unexpectedly complex, laced with rich subregularity having no bearing on the lexical derivation.

(7) accounts for the facts of morphology discussed here.

(7) **A Framework for Separation**



[3] Two alternative means of marking these derivates are unproductive: *pass-er by*, *runner-up* and *on-look-er*, *by-stand-er*).

[4] See Zwicky (1985) for "referral rules" which account for such relationships. Although Zwicky defines such rules for inflection, they clearly mark lexical derivations as well.

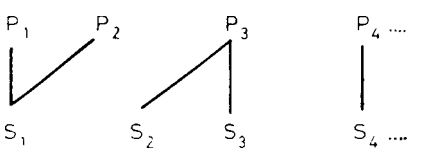
2.2. Lexeme/Morpheme-Based Morphology: Definitions. A clear implication of (7) is that (grammatical) morphemes and lexemes are radically different types of entities (Beard 1986). Unlike lexemes, grammatical morphemes and the derivational and inflectional functions they mark fall into separate closed, grammatically ordered classes. While subjective (*athlete*) and objective (*victim*) lexical classes are common, they are not related to each other via formal paradigms as are the lexical derivational subjective and objective.

(8)

	Subjective (Agentive) Adjectival	Subjective (Agentive) Nominal	Objective (Patientive) Nominal	Modalic (Instrumental) Nominal
-ee	. . .	*escap-ee*	*draft-ee*	. . .
-er	[*redd-er*]	*dry-er*	"*keep-er*"[5]	*mix-er*
-ist	*reform-ist*	*record-ist*	. . .	. . .
-ant	*depend-ent*	*migr-ant*	*rehabilit-ant*	*stimul-ant*
(-Ø)	*discriminate*	*cut-up*	*slice*	*lift*

(8) does not represent a problem of "polysemy" or "homophony". Lexical polysemy and homophony (or homonymy) are usually defined strictly in terms of the practical problems faced by lexicographers in deciding which dictionary words deserve independent entries. The only objective argument for homophony is historical, e.g. the spelling of *pair* : *pear* (homonyms) vs. that of the various meanings of *dope* (generally a polyseme). In fact nothing is gained in defining these terms over a theory of the lexicon which represents the sound-meaning relation as in (9).

(9)

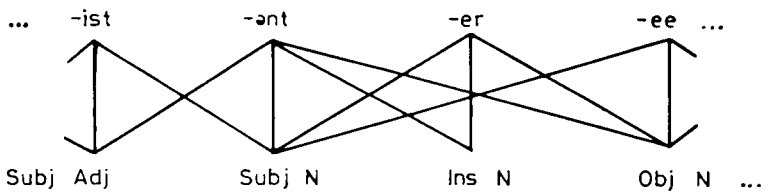


The important point of (9) is that whether one sees "homophony", "polysemy" or both here, no difficulty arises in distinguishing individual lexemes as signs in some elaboration of the term. Wierzbicka (1985, 71—73) refers to this definitional characteristic of lexemes as the "discreteness" criterion of the **Aristotelean Categorical View** of language.

[5] The admittedly marginal objective nominalizations like *fry-er, broil-er* cannot be treated as subjectives derived from middle or ergative argument structures like *The chicken fried* (*well*) unless we wish to distinguish such forms from *keep-er* (*fish*) and *loan-er* (*car*) which fail the test: *keep-er* ≠ *a fish which keeps* (*well*). (My thanks to Geert Booij for this issue.)

It is seldom possible to separate closed-class morphemes, however.

(10)



In (10) the four highly productive suffixes of (8) are listed in their complex associations with four highly productive lexical derivational categories. The interesting aspect of (10) is the impossibility of separating these morphemes into discrete signs like lexemes. Therefore, in order to maintain the discreteness criterion, we must conclude that the phonological and functional units involved are themselves discriminate and do not form discrete isomorphic associations. This entails a definition of grammatical morphemes wholly unrelated to that of lexemes.

The argument for discrete lexemes and morphemes rests on three definitional characteristics of lexemes which morphemes do not share (Beard 1986). Perhaps the most important is that lexemes do not exhibit zero phonology. Morphemes may alternate with nothing because all information necessary for semantic interpretation is in the lexeme to which it accrues or in the morpheme's paradigm or is added to the base lexeme by derivation rule. Morphemes also do not occur in phrases independently, derive, subcategorize for case and the like, for morphemes mark these relations among lexemes. Finally, morphemes are formal representations of lexical and syntactic grammar; they therefore belong to the closed classes of grammar and have no meaning at all except in the lexemes whose syntactic and lexical relations they mark. Whatever lexemes are, morphemes are not; whatever omission, reduplication, revoweling, umlaut are, so are morphemes.

We may thus partially define the lexeme (11) as a basis for a complete definition of morphemes (12).[6]

(11) A lexeme (= name) is a minimally distinguishabe (bound or free) ordered sequence of phonological segments associated with at least one sense, which may be lexically extended by lexical derivation and which belongs to an open, unparadigmatic class.

[6] The definitions of lexeme and morpheme given here are not to be confused with those of Martinet (1961). Martinet is also a structuralist and his "morphemes" and "lexemes" (variants of Martinet's "moneme") are traditional structuralist signs distinguished primarily by their belonging to closed vs. open classes. The definition of "morpheme" offered here and in Beard (1986) radically differs from that of "lexeme".

This definition covers the major class items which may fill syntactic nodes: N, V, Adj and Adv. It obviously is not sufficient in the absence of an explanation of homophony and/or polysemy (9). Adpositions are clitic or free morphemes in a class with declensional endings (Beard 1985). An "open" class may be defined here as "synchronically accessible", i.e. a class susceptible to expansion by synchronic operations and borrowing. A "closed" such class would be "synchronically inaccessible" in the same sense. "Lexical extension" here refers to derivational features like [Subject], [Object], [± Feminine], added to the sense, and not to any phonological features added to the formant of the lexeme. (See Beard (1981) for details of "lexemic extension".)

(12) A **morpheme** is the smallest phonologically distinguishable (bound or unbound) formal modification of a lexeme, which is paradigmatically ordered in a closed class, marks grammatical functions and is not susceptible to L-derivation.

A lexemic "modification" might be a free adjunct attached to a peripheral member of a phrase to mark some grammatical function of the head in the sense that *the* in *the (old) man* is a modification of *(old) man* or it may be an internal or external, bound modification as in *man:men or boy:boys*. Nothing grammatically substantive clings to the "bound-free" distinction (see also Carstairs (1981) and Marantz (to appear). Whether features like [Causative] and [Potential] are added to verbs by free-standing auxiliaries as in English or appended as affixes as in Turkish is a superficial issue unrelated to the universal issue of why these features and not others are found in languages and how they are manipulated by speakers. The crucial issue is that morphemes are associated only with internal grammatical referents and, usually, only in some context.

2.3. The Neurological Predictions of L/MBH. The L/MBH model of morphology described thus far rests on three major hypotheses about the nature of the morphological component of grammar — all in conflict with some aspect of prevailing sign based theories. First, SH implies that the nature of and conditions on rules which modify a lexeme in any manner to mark grammatical functions are independent of those which add abstract lexical or inflectional (morphosyntactic) features to lexemes. This allows the possibility of discriminate damage to syntactic, lexical and morphological processors, specifically the possibility of the presence of syntactic and/or lexical category features in the absence of morphemes marking them and vice versa, our first prediction.

This prediction is markedly at odds with the predictions of the Saussurian **Sign Base Hypothesis,** which assumes that all linguistic sound and meaning are isomorphic, biunique, form mutually implied association. Sign-based morphologies, under the Type Transparency Hypothesis, predict that the same processor which computes phonotactic sequences computes the meaning(s) or function(s) associated with them; meaning is accessible only via sound, sound only via meaning. These morphologies imply that a speaker who cannot utter the phonology should not be able to calculate its functions and those who can, should be able to.

Second, if the sound-meaning relation of the morpheme is wholly different from that of the lexeme, we would predict radically different neurological organizations for lexemes and morphemes. The distinction of lexemes and morphemes might come in the form of independent storage at different physical sites on the cerebral cortex, different locations for the access mechanisms of the parser or operationally different access mechanisms.

Discrete storage locations would support the model in (7) but would imply that the **Lexical Morphology Hypothesis** (Bloomfield 1933, Bresnan 1982, Lieber 1981, Marantz 1984), which situates lexemes and morphemes identically in the lexicon, would have to be modified or surrender any claim of type transparency. Different locations of access mechanisms or operationally distinct mechanisms would not necessarily undermine the Lexical Morphology Hypothesis if accompanied by independent evidence that morphemes and lexemes are stored in the same location. Thus we must keep in mind the distinction between storage areas and retrieval mechanisms, for this distinction underlies the crucial test of the claims of the L/MBH, on the one hand, against those of Lexical Morphology, structuralist and semiotic models, on the other.

The third notable aspect of (7), the **Integrated Morphology Hypothesis,** is supported by the work of Beard, Halle, Lieber, Marantz and others a single morphological component operates on both inflectional and lexical derivational outputs. Although the operations generating abstract derived lexemes and phrases may be discrete, for the Integrated Morphology Hypothesis to hold, the operations modifying lexemes and phrases to mark them for these operations will be conducted by the same component, or processor. Thus we should find at some level, indiscriminately affected lexical and syntactic derivational morphology in aphasia, dyslexia and normal speach error.

The competing **Split Morphology Hypothesis** (Anderson 1982, Perlmutter 1986), argues that inflectional morphology applies after all morphosyntactic adjustments have been made but that derivational morphology is an independent lexical matter accomplished in the lexicon before category selection and insertion. Again, assuming type transparency, deficits in the two morphologies should not occur as symptoms of the same syndrome but should be

affected by incidents in distinct cortical areas and should be associated with different cosymptoms if this hypothesis goes through. Moreover, since derivational morphology is assigned to the lexicon, deficits in this morphology should be associated with lexical rather than syntactic deficits.

These are not the only predictions implied by (7) and the full theory accompanying it. The fact that morphemes are operations rather than prespecified items predicts that in the performatively most transparent case morphemes will not be accessed as lexemes are. Since the definition of lexemes in this theory depends crucially on an ordered sequence of phonological segments, left-right (initial-final) ordering errors should predominate in the pathological and other speech error data. The L-rules of this theory are generative and should therefore affect lexical retrieval. How these predictions might emerge in psycholinguistic or neurolinguistic data it will not be pursued here even though the pertinent data are very suggestive.

3. The Classic Theories of Aphasia

The two major linguistic cortical processors seem to be controlled by the anterior or "Broca's" area and the posterior or "Wernicke's" area of the left cerebral cortex. Because damage to it affects the processing of "closed class", "litte" or "function" words, affixes and their functions, Broca's area is closely associated with an aphasic syndrome called "agrammatism". The central symptoms of agrammatism may be summarized as

(i) the omission or confused usage of "function words", i.e. conjunctions, prepositions, articles, pronouns, auxiliaries and copulas;

(ii) the loss or confusion of verb, noun and adjective inflection with frequent reversion to the unmarked form, e.g. verbal infinitive, nominative singular of nouns;

(iii) a reduction in the occurrence of verbs in comparison to nouns or the nominalization of verbs in some forms of agrammatic speech;

(iv) omission of arguments, e.g. subject and direct object, and misordering of syntactic constituents. Agrammatic patients exhibit little difficulty in accessing lexemes, i.e. (V,) N, Adj, Adv, but omit or find great difficulty in deploying grammatical functors. This type of speech is also called "nonfluent" because of the cooccurrence of dysprosody, the loss of control of intonation, which lends language the semblance of fluency.

Schwartz, Linebarger and Saffran (1985) provide this example of an agrammatic aphasic (M. E.) attempting to tell the story of Cinderella.

(13) Agrammatic Aphasic Speech

M. E.	*Cinderella . . . poor . . . um 'dopted her . . . scrubbed floor, um, tidy . . . poor, um . . . 'dopted . . . Si-sisters and mother . . . ball. Ball, prince um, shoe . . .*
Examiner	*Keep going.*
M. E.	*Scrubbed and uh washed and un . . tidy, uh, sisters and mother, prince, no, prince, yes. Cinderella hooked prince.* (Laughs.) *Um, um, shoes, um, twelve o'clock ball* /pInaSt/, *finished.*
Examiner	*So what happened in the end?*
M. E.	*Married.*
Examiner	*How does he find her?*
M. E.	*Um, Prince, um, happen to, um . . . Prince, and Cinderalla meet, um met um met.*
Examiner	*What happened at the ball? They didn't get married at the ball.*
M. E.	*No, um, no . . . I don't know. Shoe, um found shoe . . .*

Sensory aphasics, on the other hand, who suffer damage exclusively to Wernicke's area, become "fluent" aphasics. Sensory aphasics retain control over closed class functors and intonation, but have difficulty in recalling lexical bases. They speak in a normally intoned stream of grammatical markers, e.g. pronouns, prepositions, articles and auxiliaries, that is, in "contentless" sentences. They may recall lexical items related or unrelated to target lexemes or create nonsense neologisms (paraphasia) one form exhibits a particular difficulty in retrieving nouns (anomia). Thus while it remains a point of contention whether "control" refers to storage, retrieval, processing or some combination of these, Broca's area in some sense controls grammatical morphemes and/or the syntactic structures they mark to a significantly greater extent than Wernicke's area while the latter controls lexical bases to a greater extent than Broca's. However, as Cvetkova and Glozman (1978b) emphasize, traits of agrammatism are found to some extent in both types of aphasia.

Buckingham (1981) provides the following example of a sensory aphasic (C. B) attempting to explain a picture of a child taking a cookie as a woman spills water elsewhere in the picture.

(14) Sensory Aphasic Speech

C. B.	*Uh, well this is the . . . the* /didiŋ/ *of this. This and this and this. These things going in there like that. This is* /sen/ *things here. This one here, these two things here. And the other one here, back in this one, this one* /giʔ/ *look at this one.*

Examiner *Yeah, what's happening there?*

C. B. *I can't tell you what that is, but I know what it is, but I don't know where it is. But I don't know what's under. I know it's you couldn't say it's . . . I couldn't say what it is. I couldn't say what that is. This shu—that should be right in here. That's* /béəli/ *bad in there Anyway, this one here, and that, and that's it. This is the getting in here and that's the getting around here, and that, and that's it. This is getting in here and that's the getting around here, this one and one with this one. And this one, and that's it, isn't it? I don't know what else you'd want.*

4. Aphasiology and Current Morphological Hypotheses

4.1. Aphasiology and L/MBH. The classic theory of aphasia offers obvious support for distinct lexemes and (grammatical) morphemes. With few exceptions (discussed further on), physical locations in the cortex seem to control discrete classes of linguistic elements corresponding closely to the lexemes and morphemes of L/MBH. Moreover, the preponderance of available evidence suggests that these areas are the sites of storage and computation.

Unless we wish to consider all the linguistic areas of the brain the lexical area, it would be difficult to justify any of the Lexical Morphology Hypothesis with such evidence. This hypothesis assumes (i) that lexemes and grammatical morphemes belong to one class, the traditional isomorphic morpheme, and (ii) that both are stored in the lexicon. It can hardly be the case that the two major linguistic areas are processors but not storage sites since no variety of aphasia, short of global damage over both areas, presents a unitary symptom of loss of both classes simultaneously as identical storage location would predict. Either Wernicke's and Broca's areas in fact are areas of storage or they represent lexical and morphological retrieval devices with storage generalized throughout the brain, i.e. indistinct from general knowledge. Either way, theories which assume that morphemes and lexemes are of the same linguistic class are not substantiated in the neurological data as are those which distinguish the two.

4.2. The Separation of Derivation and Morphology. Research in aphasia in the US has most recently focussed on the relations of syntactic derivation disturbed by agrammatism. Agrammatism is traditionally explained as a deficiency of the central grammatical or syntactic processor (Bradley, Garrett and Zurif 1980). The basis for this conclusion is the disturbance or loss of morphemes which mark syntactic relations and the confusion of subject and object positions even when morphologically unmarked. The most recent research in agrammatism, however, has brought doubt on the involvement of a

unitary central grammatical processor (Berndt and Caramazza 1980, Kean 1985). Three reasons may be cited: (a) cases of agrammatism have been discovered which differ in the relative inacessibility of syntax and morphological markers, (b) the difficulty in accessing verbs, a deficit also common (and more explicable) among sensory aphasics and, finally, (c) aphasics who recognize that their sentences are ungrammatical, whose grammar thus seems independent of syntactic computation. In this section I would like to examine each of these new problems in light of the model outlined in Section 2 and show that the first two problems facing the classical theory in fact speak in favor of L/MBH while the third represents no more a problem for this hypothesis than for sign based hypotheses.

4.2.1. Syntax-Morphology Accessibility. Miceli, Mazzucchi, Menn and Goodglass (1983) discovered two Italian-speaking patients with the classic symptoms of agrammatism yet without indication of damage to any "central language processor". Both patients in this study shared classical symptoms: omissions of articles and prepositions and the substitution of infinitives for finite forms of the verb. The second patient in this study (Case 2) exhibited an overall error rate in these areas approximately twice that of the first (Case 1) thus, on the surface it seened that he was more severely agrammatic than the other. But closer examination revealed another pattern.

By computing the string length of utterances of the two patients, Miceli et al. concluded that mean string length in the speech of Case 2 was 10.1 words while that of Case 1 was only 3.6 words. Moreover, Case 2 omitted no more than three main verbs in some 70-odd clauses consisting of 600 words and his overall sample contains some 30 well-formed compound and complex sentences, ignoring morphological omissions. Miceli et al. conclude that Case 1 has a moderate syntactic deficit and a mild morphological deficit while Case 2 has an almost purely morphological deficit, suggesting discrete processors for syntax and morphology.

Grodzinsky (1984) and Grodzinsky, Swinney and Zurif (1985) have come to a similar conclusion. Working with two Hebrew-speaking agrammatic patients, Grodzinsky discovered that grammatical categories which are "strictly closed" are always morphologically marked on appropriate stem classes and are not marked only where null variants are permissible even though the proper grammatical function is seldom marked. (Carstairs 1981 defines a "strictly closed" category by the principle that one of its markers m u s t appear on the stem of the class it marks, e.g. the Latin (or Hebrew) nominal declension and verbal conjugation endings.) The relevant point which Grodzinsky makes, however, is that agrammatics misuse only appropriate affixes or revowelling schemes in such cases, never random phonological sequences. This suggests that agrammatic patients may retain a more or less intact

set of morphemes even in the absence of knowledge of the syntactic functions which they are supposed to mark.

These two studies confirm the findings of the earlier work of Tissot, Mounin and Lhermitte (1973). In an test of 19 French-speaking agrammatic aphasics, compared with 20 sensory aphasics, Tissot et al. found a significant split between those agrammatics who preserved articles and most prepositions but not the appropriate syntactic order of lexemes and those which did preserve word order but had severe difficulty with articles and prepositions. This led Tissot et al. also to divide their cases into syntactic and morphological agrammatics. Thus the growing evidence against the "classical" theory of agrammatism, showing that aphasia affects at least two subcomponents of the grammatical processor discriminately rather than damaging one indiscriminately, simply fulfills the predictions of SH.

4.2.2. Verb Accessibility. Saffran, Schwartz and Marin (1980) give the following responses from agrammatic patients attempting to describe a picture of a boy being hit in the head by a baseball.

(15) (a) *The boy is catch . . . the boy is . . . out . . . the boy is catching out.*
(15) (b) *The boy is in the . . . hits the boy.*[7]
(15) (c) *Boy is hurting to it.*
(15) (d) *Hit. The man is throwing the ball.*
(15) (e) *The boy is catch . . . the boy is hitch . . . the boy is hit the ball.*

Notice that responses (15b) and (15d—e) indicate that the problem might not be one of retrieving the correct verb so much as establishing the proper grammatical relations with the verb selected. Since the agrammatic aphasic does not have the grammatical functions s u b j e c t and o b j e c t available to him, i.e. cannot encode or decode any morphological marking of subject and object including word order (Caramazza and Zurif 1976), he resorts to a semantic strategy. His sentences will begin with a noun referring to the most prominent, usually animate, referential target. Other lexical items will be selected for their semantic relation to the first chosen item. If we assume that a modification of a lexeme includes establishing the lexeme's position in the phrase (word order), or extend our definition of "morpheme" to include modifications of phrases as well as lexemes, this data becomes yet another prediction of L/MBH.

A more serious challenge for the classic theory is the work of Miceli, Silveri, Villa and Caramazza (1984). This study elicited verbs and nouns in isolation (not in syntactic context) from agrammatic, anomic and normal

[7] Notice that patients (15a—c) seem to have access to the phonology of prepositions and particles (*out, in, to*) without control over their functions. This sort of evidence is frequently cited in the literature in support of the separation of derivation and morphology (SH).

subjects and showed that (some) agrammatics in fact have a verb deficit. Citing other evidence which shows a dissociation of verb loss and syntactic processing, however, this study concludes that verb loss results from damage "to independent cognitive systems that tend to co-occur because of the neural proximity of the two systems such that when one system is impaired there is a high probability that the other will be impaired" (Miceli et al. 1984, 218). It is still possible, therefore, to explain the difficulty of strictly agrammatic aphasics in verb selection in terms of a failure of grammatical processing rather than one of lexeme selection.

4.2.3. Judgments of Grammaticality. Linebarger, Schwartz and Saffran (1983) and Schwartz, Linebarger and Saffran (1985) have studied the ability of agrammatic patients to restart and correct themselves, behavior which suggests that they often realize that the sentences which they are speaking are agrammatic. If this is true, agrammatic aphasics might retain intact syntactic competence and only suffer a deficit of the parser. This would make it possible to argue that the lexeme/morpheme distinction is an artifact of speech production which has no correlate in competence, ergo a theory of grammar.

The study by Linebarger and her colleagues did demonstrate that their four agrammatic patients could distinguish grammatical from ungrammatical sentences at a much better than chance rate. 207 of 230 grammatical sentences were recognized as such and 163 of 221 bad sentences were recognized as such on the average. This leaves 23 of 230 grammatical sentences still misperceived as ungrammatical and 58 of 221 ungrammatical sentences mistaken as grammatical. This study begs at least two comments.

First, none of the four patients in the study responded perfectly as normal subjects presumably would. They performed worst on constructions with tag questions or reflexives, i.e. those whose interpretation rests solely on the manipulation of grammatical morphemes and relations and which gain nothing from recourse to lexical or pragmatic semantics. Thus some deficit in the basic knowledge of grammar and/or morphology clearly underlies agrammatism even in the data of Linebarger et al.

Second, recognition studies corroborating greater ability at recognizing vis à vis producing grammatical structures do not prove agrammatism to be a performative deficit if the disproportion is also present in normals. The phenomenon of comprehension quantitatively far exceeding production is a general characteristic of language performance. First and second language learners at every stage of learning are capable of comprehending far more than they are able to produce. This phenomenon probably characterizes mature speakers as well; I do not know that it has been researched or explained.

It would be reasonable to predict that the tactical ability to recognize phrases as grammatical is a subcomponent of comprehension. If so, as long as

this general phenomenon is maintained in approximately the same proportion and is not reversed or otherwise affected in agrammatic speech as compared to normal, it may not be a symptom of agrammatism. The interesting aspect of agrammatism is that, overall, the ability of a speaker to retrieve and produce grammatical material is grossly impaired in comparison to normal speakers and the proportion of impairment of comprehension and production remains the same, i.e. given the presumably normal superiority of speech comprehension over speech production. This set of facts embarrasses L/MBH no more than traditional sign-based theories.

4.2.4. Aphasiology and the Unitary Function Hypothesis. Up to this point we have seen how processors previously considered unitary in fact reflect discrete components. One prediction by L/MBH, however, predicts the integrity of components many linguists consider discrete. The Unitary Function Hypothesis predicts that the morphology of lexical derivation and inflection is unitary and identical. The Type Transparancy Hypothesis predicts that both will be computed by the same mental processor. Thus the damage to inflectional morphology witnessed in agrammatic patients should be accompanied by damage to lexical derivational morphology. The Split Morphology Hypothesis (Perlmutter 1986) predicts that lexical morphology is associated with lexical storage while inflectional morphology comprises independent postlexical operations.

Both sensory and agrammatic aphasics should exhibit lexical derivational deficits however, differences in the natures of their deficits should detectable. Agrammatic patients should have access to lexemes but have difficulty in deploying productive affixes; their speech should exhibit derivational errors and a reduction in the rate of use of productively derived lexemes but a retention of lexicalized (idiomatized) derived lexemes. Sensory aphasics, however, have difficulty accessing lexemes, the stems which are subject to lexical derivation their speech should exhibit productive affixes added to paraphasias or neologistic stems. In comprehension, we should find evidence that sensory aphasics resort more to lexical derivation in an attempt to recover the meaning of lexemes which they cannot comprehend while agrammatics ignore derivational opportunities and attempt to interpret lexical derivates as idiomatic lexical wholes.

One caveat demands attention in examining the evidence for the computation of derivational morphology. Since speakers have two means of accessing derived lexemes, storage and generation, the presence of a high percentage of properly marked derived lexemes in the speech of agrammatic speakers does not necessarily indicate split morphology. The prediction of the Unitary Function Hypothesis is that the morphological rules for marking productive lexical derivations will be affected by insult to the same area which

controls inflectional morphology. The problem here is that "productivity" is ill defined. Moreover, productivity probably varies from speaker to speaker; that is, not every speaker will productively generate the same lexical derivates. While generative lexical rules are available to speakers of languages, if they hear or use a lexical derivate often enough, they might just as well store it. Some speakers will derive *unaccountability* others, who use it more, will store it.

Evidence for the Unitary Function Hypothesis is found in the work of Tissot, Mounin and Lhermitte (1973) and Dressler (1977, 1979) but most of all in that of Luria's students, Akhutina, Cvetkova, Glozman and Teplickaia. Glozman (1974), Cvetkova and Teplickaia (1975) and Cvetkova and Glozman (1978a, 1978b) found that all aphasics exhibit agrammatic speech but those diagnosed as suffering from motor (Broca's) aphasia cannot account for the compositionality of derived lexemes and treat them as units. Patients suffering from posterior aphasias, acoustico-mnestic (sensory) and semantic in Luria's terms, rely heavily on compositional elements and tend to treat them transparently, ignoring lexical idiomaticities. Dressler (1977) found similar evidence in German-speaking patients, concluding that the "lexicon and word formation seem to be different components stored in different parts of the brain" and that the latter is a part of the same component which controls syntax.

Cvetkova and Glozman (1978a) conducted two experiments with motor and dynamic aphasics, suffering from anterior insult, and semantic and acoustico-mnestic aphasics, suffering from posterior insult. The acoustico-mnestic (sensory) aphasics exhibit a significantly lower rate of agrammatism than the other three types and have all the general symptoms of sensory aphasia. Semantic aphasics share the symptoms of Broca's and sensory aphasias. In the Cvetkova and Glozman experiments, all categories of aphasics were compared with a normal control group. The first test involved the analysis of sentences without spaces, e.g.

(16) (a) *nakruglomstolestoitvazasosennimicvetami*
'onacircular tablestandsavasewithfallflowers'

(b) *Na krug-l + om stol + e stoi + t vaz + a s osen-n + imi cvet + ami*
on circul-ar+Loc table+Loc stand+s vase+Nom with autumn-a+Ins
Sg Sg Sg Pl
flower+Ins
Pl
'On a circular table stands a vase with autumn(al) flowers'

Cvetkova and Glozman discovered that the strongly agrammatic motor and dynamic (also semantic) aphasics, on the one hand, and the marginally agrammatic sensory patients, on the other, tend to make different kinds of mistakes on this test. One class of mistakes were of the type *nakruglom* 'onthe-

circular', *istala* 'andbecame', *jabyl* 'Iwas'. Cvetkova and Glozman calls these "lexical" errors because they overlook grammatical markers (prepositions like *na* 'on', conjunctions like *i* 'and', pronouns like *ja* 'I') and treat phrases as unitary lexical items, according to their semantic content. 91% of the motor aphasics' errors on the first test were of this type as compared to 18% of the sensory aphasics'.

The sensory aphasics tended to make what Cvetkova and Glozman consider "derivational" errors for they indicate an erroneous lexeme division at morpheme boundaries, e.g. *krug lom, stal a.* This suggests that the sensory aphasics overgeneralize morphological analysis, perhaps not recognizing stems. 82% of the errors of the sensory aphasics were of this type while this type accounted for only 9% of the motor aphasics' errors. While the motor aphasics, presumably with the most greatly reduced morphological capacity, did not mistake affixes for lexemes, they did tend to fail to distinguish grammatical morphemes from the lexemes to which they attach or with which they form a grammatical unit. The sensory aphasics, however, presumably with a greatly impaired ability for lexical retrieval but minimally impaired access to grammatical rules, tended to isolate the lexemes by separating all grammatical morphemes from them. This overseparation is always along correct morphological lines with indifference to lexical (*krug lom*) and inflectional (*stal a*) morphology.

Cvetkova and Glozman's second test was a classification test consisting of words written on 15 cards which normal and aphasics with the same four syndromes as defined by Luria, were asked to classify according to their semantic relationships. The cards contained genuine lexical derivates (*let-at'* 'to fly' and *let-atel'nj* 'flying (Adj)', *veter* 'wind' and *vetr-jannyj* 'wind(y)') and potential but false derivates (*vert-o-let* 'revolve-o-fly' = 'helicopter' and *vertikal'* 'vertical (line)', *rubl'* 'ruble' and *rub-it'* 'chop').

Again two types of errors were observed again corresponding to the type of general impairment; the motor (dynamic and semantic) as opposed to sensory aphasics. The motor aphasics tended to coclassify any lexemes with formally similar stems, reading from left to right, e.g. *vertolet* : *vertikal'*, *rubin* 'ruby' : *rubit'*, significantly more than did the sensory aphasics (66% vs. 34% of their errors) and to ignore genuine lexical derivational relationships like *leta-t'* : *leta-tel'nyj*. Again, this was taken as an indication of overgeneralized lexical analysis at the expense of morphological (derivational).

The sensory aphasics' performance was just the opposite, only 20% of their errors were lexical in this sense while 80% were morphological (derivatio nal). The sensory aphasics, for example, tended to co-classify words like *namererenie* 'intention' : *merit'* 'measure', *rublennyj* 'chopped' : *rubl'* 'ruble', where genuine morphological boundaries are detectable though synchronically irrelevant (idiomatized). The sensory aphasics tended to "reetymologize" these

words, i.e. analyze them as derivationally related; they confused phonologically similar but semantically unrelated stems without morphological boundaires, e.g. *rubin*: *rubit'*, significantly less.

5. Conclusion

The L/MBH makes three major predictions: (a) there are two basic units of grammar, the "lexeme" and "morpheme" as described here, (b) lexical and syntactic derivation are independent of all rules of morphological marking and (c) all morphology is unitary and integrated. The simplest imaginable versions of the Sign Base Hypothesis, the Lexical Morphology Hypothesis, claims that all morphology is lexical. This hypothesis makes at least two predictions as to basic lexical units and processes, (a) all units are lexical signs and (b) only one process handles them: lexical insertion. Thus the Lexical Morphology Hypothesis is far simpler and more constrained than the L/MBH, shifting the burden of proof upon the latter.

But all the fundamental predictions of L/MBH are borne out by neurological evidence. Moreover, the data of language acquisition and normal speech errors add further confirmation. Garrett (1980), in summing up his extensive research in speech error analysis, concludes " . . . that shift errors [e.g. *I haven't satten down and writ . . .*; *I want to eated.*; *I had forgot abouten that*; *. . . point outen*] attach morphemes to stems without regard to any factor other than word boundaries and word form — in particular, without regard to the lexical identity or even grammatical category of the error site." His findings confirm similar evidence discussed by Bierwisch (1981) for SH, that the phonological formants of affixes are independent of their functions (*outen* cannot be in the past tense for prepositions have no tense). Moreover, he argues that shift errors like these must occur at a processing level other than that which accounts for "word-exchange" errors, e.g. *I don't know that I would hear one if I knew it*). Garrett, like Bierwisch, points out that the word-exchange errors generally reflect lexically marked morphology as this example does (*knew*), while morpheme-shift errors apparently do not (crucial data is scarce). Errors at the point of lexical insertion and morphological realization in (7) account for Bierwisch's and Garrett's analyses of the speech error data neatly.

Evidence of the Separation Hypothesis is also found in language acquisition studies. Several studies have shown that children learn the cases and their functions (case "semantics") prior to mastery of affixational rules, e.g. Bogoyavlenskii (1973) and the sources she cites. This disjuncture is similar to the overgeneralization of productive affixes which characterize certain stages of acquisition. During these stages, children who previously used, for example, irregular and subregular past tense verb forms such as *caught*, *hit* and *saw*

correctly, begin saying *catched, hitted* and *seed* in alternation with the grammatical, subregular forms. However, as Maratsos (1979, 312—313) points out, since both alternates are used in the appropriate grammatical situations, the use suggests a disjuncture between knowledge of form and that of the function.

This same evidence is incompatible with the LAH though not with Word-and-Paradigm models which separate lexical from inflectional derivations (Matthews 1972, Anderson 1982, 1984). However, word-and-paradigm models will probably have to adjust to the evidence of or the Integral Morphology Hypothesis and combine inflectional and lexical morphology into one component. The second prediction not only survives the neurological evidence but finds support there as a secondary characteristic of linguistic behavior. The prediction that knowledge of the functional constraints on lexical as well as inflectional derivation are independent of the affixational constraints contradicts the predictions of all sign-based morphologies, Word-and-Paradigm included. All sign-based models will require heavy translational machinery in performance theory. If we accept the Type Transparency Hypothesis, therefore, Lexeme-Morpheme Base models of morphology are decidedly superior (more isomorphic with reasonable performance models) to Sign Based models.

References

Akhutina, T. V.: Neĭrolingvisticheskiĭ analiz dinamicheskoĭ afazii. Moscow, Moscow University Press 1975.

Anderson, St.: Where's morphology. In: Linguistic Inquiry 13 (1982), 571—612.

Beard, R.: A semantically based model of a generative lexical wordformation rule for Russian adjectives. In: Language 52 (1976), 108—120.

Beard, R.: The Indo-European lexicon: a full synchronic theory. North-Holland Linguistic Series 44, North-Holland, Amsterdam 1981.

Beard, R.: Prepositions and case. In: Zbornik Matice srpske za filologiju i lingvistiku 27—28 (1984—1985), 57—63.

Beard, R.: On the separation of derivation from morphology: toward a lexeme/morpheme-based morphology. Bloomington, Indiana University Linguistics Club 1986.

Berndt, R. and Caramazza, A.: A redefinition of the syndrome of Broca's aphasia. Implications for a neuropsychological model of language. In: Applied Psycholinguistics 1 (1980).

Berndt, R. and Caramazza, A.: Syntactic aspects of aphasia. In: Sarno, M. (ed.): Acquired Aphasia. Academic Press, New York 1981.

Berwick, R. and Weinberg, A.: The role of grammars in models of language use. In: Cognition 13 (1983), 1—61.

Berwick, R. and Weinberg, A.: The grammatical basis of linguistic performance: Language use and acquisition. MIT Press, Cambridge, Mass. 1984.

Bierwisch, M.: Linguistics and language error. In: Linguistics 19 (1981), 581—626.

Bloomfield, L.: Language. Henry Holt and Company, New York 1933.

Bogojavlenskiĭ, D. N.: The acquisition of Russian inflections. In: Ferguson, Ch. & Slobin, D. (eds): Studies of Child Language Development. Holt, Rinehart and Winston, New York 1973.

Bresnan, J.: A realistic transformational grammar. In: Halle, M., Bresnan, J. and Miller, G. (eds): Linguistic Theory and Psychological Reality. MIT Press, Cambridge, Mass. 1978.

Bresnan, J. (ed.): The mental representation of grammatical relations. MIT Press, Cambridge, Mass. 1982.

Bradley, D., Garrett, M. and Zurif, E.: Syntactic deficits in Broca's aphasia. In: Caplan, D. (ed.): Biological Studies of Mental Processes. MIT Press, Cambridge, Mass. 1980, 269—286.
Buckingham, H. W., Jr.: Where do neologisms come from? In: Brown, J. (ed.): Jargonophasia. Academic Press, New York 1981.
Carrier-Duncan, J.: Linking of thematic roles in derivational word formation. In: Linguistic Inquiry 16 (1985), 1—34.
Carstairs, A.: Notes on affixes, clitics and paradigms. Indiana University Linguistics Club, Bloomington 1981.
Chomsky, N.: Lectures on Government and Binding. Foris, Dordrecht 1981.
Cvetkova, L. and Teplickaia, S. V.: Rol' slovoobrazovaniia v formirovanii parafazii u bol'nykh s afazieĭ. In: Materialy u vsesojuznogo simpoziuma po psikholingvistike i teorii kommunikacii. AN SSSR 1975, 179—183.
Cvetkova, L. and Glozman, Zh.: Neirolingvisticheskii analiz nekotorykh aspektov morfologii slova. In: Voprosy psikhologii. No. 1 (1978a), 91—99.
Cvetkova, L.: and Glozman, Zh.: Agrammatism pri afazii. Moscow University Press, Moscow 1978b.
Cvetkova, L. and Glozman, Zh.: O vzaimootnoshenii psikhologicheskogo i grammaticheskogo aspektov rechi pri afazii. In: Voprosy psikhologii. No. 6.
Dressler, W. V.: Elements of a polycentristic theory of word-formation. In: Wiener Linguistische Gazette 15 (1977), 13—32.
Dressler, W. U.: On a polycentristic theory of word-formation. In: Dressler, W. U. & Meid (eds): Proceedings of the XIIth International Congress of Linguists. Institut für Sprachwissenschaft, Innsbruck 1979.
Dressler, W. U.: On the predictiveness of Natural Morphology. In: Journal of Linguistics 21 (1985), 321—337.
Dressler, W. U. and Barbaresi, L.: How to fix interfixes: on the structure and pragmatics of Italian (and Spanish, Russian, Polish) antesuffixal interfixes and of English intermorphic elements. In: this volume, 53—67.
Fodor, J.: The Modularity of Mind; An Essay on Faculty Psychology. MIT Press, Cambridge, Mass. 1983.
Fodor, J., Bever, Th. and Garrett, M.: The Psychology of Language an Introduction to Psycholinguistics and Generative Grammar. McGraw-Hill, New York 1974.
Garrett, M.: The limits of accomodation: arguments for independent processing levels in sentence production. In: Fromkin, V. (ed.): Errors in linguistic performance; slips of the tongue, ear, pen, and hand. Academic Press, New York 1980.
Glozman, Zh.: K voprosu o narushenii morfologicheskoi struktury rechi pri afaziĭ. In: Voprosy psikhologii, No. 5 (1974), 81—87.
Grodzinsky, Y.: The syntactic characterization of agrammatism. In: Cognition 16 (1984), 99—120.
Grodzinsky, Y., Swinney, D. and Zurif, E.: Agrammatism: structural deficits and antecedent procession disruptions. In: Kean, M.-L. (ed.), 65—81.
Kean, M.-L. (ed.): Agrammatism. Academic Press, Orlando, Florida 1985.
Klavans, J.: The independence of syntax and phonology in cliticization. In: Language 61 (1985), 95—120.
Lieber, R.: On the Organization of the Lexicon. Indiana University Linguistics Club, Bloomington, Indiana 1981.
Linebarger, M., Schwartz, M. and Saffran, E.: Sensitivity to grammatical structure in so-called agrammatic aphasics. In: Cognition 13 (1983), 361—392.
Marantz, A.: On the nature of grammatical relations. In: Linguistic Inquiry Monographs 10. MIT Press, Cambridge, Mass. 1984.
Maratsos, M.: How to get from words to sentences. In: Aaronson, D. & Rieber, R. (eds): Psycholinguistic Research: Implications and Applications. Lawrence Erlbaum Associates, Hillsdale, N. J. 1979.
Marchand, H.: The categories and types of present-day English wordformation. 2nd ed. C. H. Beck, Munich 1969.
Marcus, M.: A theory of syntactic recognition for natural language. MIT Press, Cambridge, Mass. 1980.
Martinet, A.: Éléments de linguistique générale. Colin éditions, Paris, 1961.
Matthews, P. H.: Inflectional Morphology: A Theoretical Study Based on the Aspects of Latin Verb Conjugation. Cambridge University Press, Cambridge 1972.
Mayerthaler, W.: Morphologische Natürlichkeit. In: Linguistische Forschungen 28. Athenaion, Wiesbaden 1981.

Miceli, G., Silveri, M. C., Villa, G. and Caramazza, A.: On the basis for the agrammatic's difficulty in producing main verbs. In: Cortex 20 (1984), 207–220.
Miceli, G., Mazzucchi, A., Menn, L. and Goodglass, H.: Contrasting cases of Italian agrammatic aphasia without comprehension disorder. In: Brain and Language 19 (1983), 65–97.
Perlmutter, D.: The split morphology hypothesis: evidence from Yiddish. Paper presented at the Milwaukee Morphology Meeting, 15th Annual University Wisconsin-Milwaukee Linguistics Symposium, Milwaukee, April 4–6, 1986.
Pesetsky, D.: Morphology and logical form. In: Linguistic Inquiry 16 (1985), 193–246.
Pounder, A.: The semantic organization of word-formation paradigms in diachrony. In: this volume, 181–195.
Sadock, J.: Autolexical syntax: A proposal for the Treatment of noun incorporation and similar phenomena. In: Natural Language and Linguistic Theory 3 (1985), 379–439.
Saffran, E., Schwartz, E. M. and Marin, O.: Evidence from aphasia: Tsol components of production. In: Butterworth, B. (ed.): Speech and Talk, Volume I, Academic Press, New York 1980.
Schwartz, M., Linebarger, L. and Saffran, E.: The status of the syntactic deficit theory of agrammatism. In: Kean, M.-L. (ed.), 83–124.
Szymanek, B.: Morphological and Phonological Constraints on Suffixation; Derivation of Polish and English Adjectives. Catholic University Press, Lublin 1985.
Tissot, R., Mounin G. and Lhermitte, F.: L'agrammatisme. Charles Dessart, Brussels 1973.
Wierzbicka, A.: Lexicography and Conceptual Analysis. Karoma, Ann Arbor 1985.
Wurzel, W.: Flexionsmorphologie und Natürlichkeit. In: Studia Grammatica. Akademie der Wissenschaften der DDR, Berlin 1984.
Zwicky, A.: Rules of allomorphy and phonology-syntax interactions. In: Journal of Linguistics 21 (1985), 431–436.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 25–37 (1986)

THE ROLE OF CONCEPTUALIZATION RULES IN THE INTERPRETATION OF MORPHOLOGICALLY COMPLEX WORDS

RUDOLF P. BOTHA

(University of Stellenbosch, South Africa)

1. Introduction*

Allow me to start with a statement of metascientific belief: I take the pursuit of depth of insight or understanding rather than gross coverage of data to be a legitimate aim of linguistic inquiry. Many of you would have little trouble in accepting this aim for the study of word-formation or morphology too. Many of you, likewise, would agree, without much reservation, that in order to acquire the desired understanding or insight, we need principled theories of considerable deductive depth and unifying power. What I would like to do now, is to explore some of the consequences of this metascientific position for an analysis of the semantics of Afrikaans reduplication.[1] In particular, I will argue that this position drives us towards recognizing the existence of a class of "semantic" devices which, for want of a better name, I will call c o n c e p t u a l i z a t i o n r u l e s. I have presented this argument elsewhere in much more detail, but it would be presumptious to assume that you are familiar with it.[2]

On conventional analyses, the so-called meanings said to be expressed by reduplication in Afrikaans are stunningly diverse. These include

(A) considerable number ("many R's") as in

(1) *Die kinders drink b o t t e l s - b o t t e l s limonade.*
the children drink bottles bottles lemonade
'The children drink bottles and bottles of lemonade.'

(2) *B a k k e - b a k k e veldblomme versier die tafels.*
bowls bowls veld flowers decorate the tables
'The tables are decorated with wild flowers by the bowlful.'
[I use "R" in the paraphrases to represent the referent to which the base (α) refers via its meaning (A).]

* This paper was written for oral presentation at The International Conference on Word Formation held on 1—4 May 1986 at Veszprém, Hungary. To allow for this mode of presentation, the style of the paper was accordingly kept relatively informal where the argument did not require it to be otherwise.

[1] This position forms part of the so-called (lax) Galilean Style of inquiry, as is clear from Botha 1982, 1984.

[2] Cf. Botha 1984, 1986.

(B) limited number ("some R's") as in

(3) *Die pad was e n t-e n t sleg.*
the road was stretch stretch bad
'The road was bad in some (scattered) stretches.'

(4) *Jan vergeetsy vrees r u k-r u k.*
John forget his fear time time
'Occasionally John forgets about his fear for a while.'

(C) distribution ("scattered R's") as in

(5) *Die skape wei t r o p p e - t r o p p e op die vlakte.*
the sheep graze flocks flocks on the plain
'The sheep are grazing on the plain in several scattered flocks.'

(6) *Die gras het k o l - k o l verdroog.*
the grass has patch patch withered
'The grass has withered in (some) scattered patches.'

(D) serial ordering ("the one R after the other") as in

(7) *Hy krap die verf l a a g - l a a g af.*
he scrapes the paint layer layer off
'He scrapes the paint off in one layer after another.'

(8) *Die studente skryf die eksamen s t u k - s t u k.*
the students write the exam piece piece
'The students write the exam in instalments.'

(E) collectivity ("in sets/groups of R") as in

(9) *Hy dra t i e n - t i e n boeke die trap op.*
he carry ten ten books the stairs up
'He carries the books up the stairs in one ten-book batch after another.'

(10) *Susan sluk die pille d r i e - d r i e in.*
Susan swallow the pills three three in
'Susan swallows the pills in sets of three.'

(F) iteration ("to R more than once/repeatedly") as in

(11) *Hy l e k - l e k oor sy droe lippe.*
he lick lick over his dry lips
'He licks and relicks his dry lips.'

(12) *Sy kop k n i k - k n i k van vermoeienis.*
his head nod nod of weariness
'His head repeatedly nods with weariness.'

(G) *continuation* ("to R continuously/for some time") as in

(13) *Die donder r a m m e l - r a m m e l in die verte.*
the thunder rumble rumble in the distance
'A continual rumble of thunder may be heard.'

(14) *Die bedelaar d r e n t e l - d r e n t e l doelloos in die park rond.*
the beggar saunter saunter aimlessly in the park about
'The beggar has been sauntering aimlessly in the park for some time.'

(H) attenuation ("to R (more than once) tentatively/hesitantly/non-intensely") as in

(15) *Die dokter v a t - v a t aan die swelsel.*
the doctor touch touch on the swelling
'The doctor tentatively feels the swelling a couple of times.'

(16) *Hy s k o p - s k o p teen die deur.*
he kick kick against the door
'He gives the door a few exploratory kicks.'

(I) intensity ("very R") as in

(17) *Hulle eet d i k - d i k snye brood.*
they eat thick thick slices bread
'They eat thumping thick slices of bread.'

(18) *Sy het a m p e r - a m p e r haar been gebreek.*
she has nearly nearly her leg broken
'She very nearly broke her leg.'

(J) emphasis ("emphatically/specifically/definitely/just R") as in

(19) *Die ongeluk het h i e r - h i e r gebeur.*
the accident has here here happened
'The accident happened right there/on this very spot.'

(20) *Hulle doen die werk s a a m - s a a m.*
they do the work together together
'They do the work very much as a team effort.'

The ten meanings A—J by no means exhaust the list that has conventionally been claimed to be expressed by reduplication in Afrikaans. Others I have dealt with elsewhere too.

For a linguist who is after insight or understanding, it is rather puzzling how such a diversity of meanings could be expressed in one language by the single simple formal means of total reduplication. Clearly, no insight could be gained by means of an analysis that postulated a separate semantic interpretation rule for each of the ten or more distinct meanings associated with Afrikaans reduplications. An analysis which reduced the diversity and com-

plexity of these meanings to a single, unifying interpretation rule would obviously make a more significant contribution to our understanding of the semantics of these reduplications. And I will claim that all Afrikaans reduplications are, as a matter of fact, subject to a single, simple semantic interpretation rule, namely (21).

(21) Interpret $[\alpha_i\ \alpha_i]$ as [A INCREASED]
(where A represents the sense or meaning of α and INCREASED represents an abstract semantic unit).

Less formally, this rule may be understood as saying that by reduplicating a base form, the information is conveyed that the entity or entities referred to by the base form is taken to be increased in some dimension.

The central question to be answered, is how so simple a rule is able to account for the diversity and specifity of the meanings conventionally attributed to Afrikaans reduplications. The essence of my answer to this question is that the diversity and specifity of these meanings are a function of the interaction between the interpretation rule (21) and general semantic and/or conceptual devices independent of it. Central amongst the latter devices are the conceptualization rules that I have mentioned earlier.

Within the restricted scope of a single paper I cannot hope to reanalyze from this perspective all the meanings claimed to be expressed by Afrikaans reduplications. I can do little more than give a number of illustrative reanalyses — specifically of the so-called "verbal" meanings of "iteration", "continuation" and "attenuation". Before looking at "iteration", I have to draw your attention to a terminological practice that I will adopt: expressions such as "meaning", "semantic content", "information content", "unit of semantic or information content" and "semantic reading" will be used informally as synonyms. Nothing in my analysis hinges on this terminological variation

2. Iteration ("to R repeatedly")

Let us start with "iteration" and consider the make-up of the meaning or total information content associated with verb reduplications such as *lek-lek* in (11) and *knik-knik* in (12).

(a) The lexical base (*lek*, *knik*) contributes two units of meaning to the total semantic content of these reduplications. The first unit distinguishes the meaning of, for example, *lek* from that of *knik*, and other nonsynonymous lexical items. These units of meaning of *lek* and *knik* may be represented schematically as (22a) and (b) respectively.

(22) (a) [LICK][3]
(b) [NOD]

The second is a unit shared by the meaning of *lek* and the meaning of *knik*. This unit of meaning may be represented as in (23).

(23) [TEMPORAL ACT/EVENT][4]

(b) The interpretation rule (21) contributes the unit of meaning [INCREASED] to the total information content of the reduplications under consideration, giving the composite readings of (24).

(24) (a) [LICK, TEMPORAL ACT/EVENT, INCREASED]
(b) [NOD, TEMPORAL ACT/EVENT, INCREASED]

(c) At this point the question arises as to how the semantic specification/unit [INCREASED] is to be understood or conceptualized in conjunction with the meanings of the lexical bases *lek* and *knik* respectively. And this is where what I call conceptualization rules enter into the analysis. A conceptualization rule specifies how [INCREASED] has to be conceptualized in conjunction with the meanings of *lek* and *knik* respectively. Since the meaning of both *lek* and *knik* incorporates the semantic unit [TEMPORAL ACT/EVENT], a conceptualization rule with the gist of (25) applies to the readings (24a) and (b).

(25) Conceptualize [INCREASED] as [INCREASED IN TIME] if it occurs in conjunction with the semantic unit [TEMPORAL ACT EVENT].[5]

Given this rule, the composite readings (26a) and (b) may be formed.

(26) (a) [LICK, TEMPORAL ACT/EVENT, INCREASED IN TIME]
(b) [NOD, TEMPORAL ACT/EVENT, INCREASED IN TIME]

(d) In the composite readings (26a) and (b), however, the unit [INCREASED IN TIME] is only partially amalgamated with the semantic units

[3] Capitals are used to represent abstract semantic units. These units are not necessarily minimal but may be further decomposable.
[4] On Lyons's (1977, 484) analysis an act is an event that is under control of an agent.
[5] This rule, most likely, will turn out to represent a subcase of a more general rule applying to reduplications based on nouns, adjectives, and adverbs too.

[LICK] and [NOD]. The reason for this is that licking and nodding represent a particular kind of temporal act or event: in the terminology of Jackendoff (1983, 246) they are temporally bounded events or acts. This feature of the (projected) referent of the verbs *lek* and *knik* may be represented in their meaning by the semantic unit [BOUNDED] which constitutes a unit of so-called aspectual meaning.[6] The representations (26a) and (b) thus have to be refined to read as (27a) and (b) respectively.

(27) (a) [LICK, TEMPORAL ACT/EVENT, BOUNDED, INCREASED IN TIME]
(b) [NOD, TEMPORAL ACT/EVENT, BOUNDED, INCREASED IN TIME]

To amalgamate the semantic unit [INCREASED IN TIME] with [BOUNDED] a conceptualization rule with the content of (28) is required:

(28) Conceptualize the unit of content [INCREASED IN TIME] as [ITERATED] if it occurs in combination with the unit of aspectual meaning [BOUNDED].

Clearly, a bounded event can occur for an increased time only if it is conceptualized as being repeated more than once. Applied to the readings specified in (27a) and (b), the conceptualization rule (28) forms the readings (29a) and (b) respectively.

[6] As has been noted by, for example, Holisky (1981, 28) "The term 'aspect' has almost as many definitions as there are linguists who have used it . . .". Platzack (1979, 39) draws a distinction between aspect and aktionsart:

> "Whereas aktionsart has to do with the inherent temporal constitution of a situation, independent of deictic time (i.e., time in its relation to speaker and hearer), we will use the term *aspect* to refer to the way a speaker (or writer) chooses to present a situation in relation to deictic time, provided that the language offers a systematic way to express the choice in question. Thus, aspect is intimately connected to the use of a sentence (or, as we will prefer to say, to the possible use of a sentence). To describe the aktionsart referred to by a sentence, we do not have to take into consideration how the sentence is related to the communicative situation (though such a relation may be taken into consideration when we like to disambiguate a sentence in cases where a given string of words is able to refer to several aktionsarten). However, in order to describe the aspect of a sentence, this relation is of utmost importance."

I will use the term aspect, and a derived form such as aspectual, to denote Platzack's aktionsart. This is a common usage of the term aspect, as is clear from Comrie's (1976, 3) discussion in which he presents "a general definition of aspect" according to which "aspects are different ways of viewing the internal temporal constituency of a situation", a definition attributed by him to Holt (1943, b). For remarks on the history of the aspectual notion of "boundedness" cf., e.g., Platzack 1979, 70ff., and Dahl 1981, 79–81. For critical comments on Platzack's distinction between aspect and aktionsart cf. Andersson 1984, 200ff.

(29) (a) [LICK, TEMPORAL ACT/EVENT, ITERATED]
(b) [NOD, TEMPORAL ACT/EVENT, ITERATED]

As a unit of meaning, then, [ITERATED] need not be specified directly by an interpretation rule such as (21). It is a derived unit, established through the interaction of this rule with the conceptualization rules (25) and (28). Note that the extension of the parameter of boundedness from things to acts/events, which is crucial to this analysis, has been argued for on independent grounds.[7]

3. Continuation ("to R continuously")

Let us next consider the meaning of "continuation" expressed by Afrikaans reduplications. The total information content associated with verb reduplications such as *rammel-rammel* in (13a) and *drentel-drentel* in (13b) is parallel, in composition, to that of iterative reduplications such as *lek-lek* and *knik-knik*. The difference between iteration and continuation reduces to a difference in aspectual meaning between *lek-lek* and *knik-knik* on the one hand and *rammel-rammel* and *drentel-drentel* on the other hand. The base verbs of the former type have the unit of aspectual meaning [BOUNDED]; the base verbs of the latter have the unit of aspectual meaning that Jackendoff (1983, 246) calls [UNBOUNDED]. This implies that the unit of aspectual meaning [UNBOUNDED] has to be incorporated in the composite readings (30a) and (b) for *rammel-rammel* and *drentel-drentel* respectively:

(30) (a) [RUMBLE, TEMPORAL ACT/EVENT, UNBOUNDED, INCREASED IN TIME]
(b) [SAUNTER, TEMPORAL ACT/EVENT, UNBOUNDED, INCREASED IN TIME]

For the further conceptual amalgamation or integration of the readings (30a) and (b) we require a conceptualization rule with the purport of (31).

[7] As has been noted by a number of linguists, most recently by Jackendoff (1983), there is a parallelism between iterating events and increasing things in number. At a linguistic level, therefore, "iterativity" and "plurality" are fundamentally similar notions. Platzack (1979, 79ff.) and others use the semantic feature "+/—DIVID" as a semantic correlate of Chomsky's (1965) syntactic feature "+/—COUNT". Following Teleman (1969), Platzack (1979, 81) argues that the feature "DIVID" is useful for capturing the "count-mass" distinction in the description of (Swedish) noun phrases. Teleman suggests that this feature can be used for the description of verbs too, with "—DIVID" to "durative" verbs. However, Platzack argues that the feature should not be assigned to the verb, but to the sentence, because it is the situation corresponding to the sentence that should be described in terms of a k t i o n s a r t e n. For a further illustration of the explanatory power of the notion of "boundedness" cf. Heinämäki 1984, 155 ff.

(31) Conceptualize the unit of content [INCREASED IN TIME] as [CONTINUED] if it occurs in combination with the unit of meaning [UNBOUNDED].

What this rule says, in essence, is that by increasing an unbounded temporal act or event one gets a single extended act or event of the same sort. Applied to the composite readings under consideration, rule (31) gives the conceptually more highly developed readings of (32).

(32) (a) [RUMBLE, TEMPORAL ACT/EVENT, CONTINUED]
(b) [SAUNTER, TEMPORAL ACT/EVENT, CONTINUED]

The distinction between iteration and continuation, therefore, reflects an aspectual difference between the base verbs of reduplications. This distinction is acted on by different conceptualization rules or different subcases of the same, more general, conceptualization rule. In sum: the distinction between iteration and continuation clearly need not be accounted for directly by the interpretation rule (21).[8]

4. Attenuation ("to R more than once non-intensely")

This brings us to the meaning of "attenuation" expressed by Afrikaans reduplications. The total information content of reduplications such as *vat-vat* in (15) and *skop-skop* in (16) incorporates what appears to be a mysterious component. This component, which has conventionally been characterized as "tentatively/hesitantly/non-intensely", will be represented below by the abstract specification [ATTENUATED]. The question is how this unit can be a component of a composite content to which the interpretation rule (21) contributes the semantic unit [INCREASED]. As part of the total information content of verb-based reduplications, one would expect the latter unit to be conceptualized on an intensity scale as "more intensely" rather than "less intensely", "tentatively", etc. Closer analysis shows, however, that there is in fact nothing mysterious about the way in which the unit [ATTENUATED] is derived as a component of the information content of reduplications such as *vat-vat* and *skop-skop*.

(a) The bases (*vat, skop*) of such reduplications are verbs that have the unit of aspectual meaning [BOUNDED]. The reduplications, consequently,

[8] In a study of Swedish aktionsarten, Platzack (1979, 124ff.) too argues that "iterativity" and "durativity" are not fundamental notions. Bridgen (1984) similarly argues that in Finnish iterativity and habituality are not basic aspectual notions.

are assigned an interative reading in the way desribed with reference to *lek-lek* and *knik-knik* in (d) of §2 above.

(b) The unit of content [ATTENUATED] represents another derived component of the total information content of verb-based reduplications such as *vat-vat, skop-skop*, etc. Let us consider the sentences (33) and (34) to get a better grasp of the nature of this unit of meaning.

(33) *Hy* skop-skop *teen die deur.*
he kick kick against the door
'He tentatively kicks the door a couple of times.'

(34) *Hy* sluit-sluit *die deur.*
he lock lock the door
'He tentatively locks the door a couple of times.'

The total information content of *skop-skop* in (33) includes both the components [ITERATED] (expressed by "a couple of times" in the paraphrase) and [ATTENUATED] (expressed by "tentatively" in the paraphrase). Native speakers of Afrikaans intuitively judge this sentence "to make sense", etc. If sentence (34) is interpreted in a parallel way, however, speakers judge this sentence to be "nonsensical", "illogical", etc. This difference in acceptability between the two sentences may be explained indirectly with reference to the nature of the events or acts denoted by *skop* and *sluit* respectively. Note that the event/act denoted by *sluit* has a certain conclusiveness or finality. The event/act denoted by *skop*, by contrast, lacks this feature: it is inconclusive or non-final. Obviously, it is impossible to repeat an event/act that has this property of finality in a relatively short time-span. And this is why sentence (33) is "nonsensical" to speakers of Afrikaans. To put it differently, the finality of the event/act of locking something precludes the possibility of its occurring repeatedly within the same short time-span, without the intervention of another act/event, specifically an "unlocking" event/act. In the case of *skop*, by contrast, the event/act lacks this finality. Consequently, it may be repeated within a relatively short time-span. For this reason native speakers have no problem in "making sense" of sentence (33). The essence of the semantic difference between reduplications such as *sluit-sluit* and reduplications such as *skop-skop* may, therefore, be captured by the generalization (35).

(35) If an event/act has the property of finality, it cannot occur/be performed more than once in a relatively short time-span.

Evidently, events/acts that have the property of finality cannot occur/be performed less intensely. That is, such events/acts cannot be attenuated.

The repeatability of an act/event, thus, indicates its attenuability. The repetition of an event may therefore be interpreted as an indication of its being attenuated on a scale of intensity. Notice that what we have here is indication, not logical entailment.

The question, of course, is how the difference between the events/acts denoted by *sluit* and those denoted by *skop* may be expressed by a conceptualization rule operating on the semantic units composing the meanings [LOCK] and [KICK]. Finding an answer to this question is a matter of determining whether the difference can be accounted for in aspectual terms. Note that verbs such as *sluit* denote events/acts that have been called "achievements" by Vendler (1967, 103). On his view, an achievement — e.g., to arrive at a destination, to win a race, to reach the top of a hill, to forget or remember something — is an event or act that occurs at a single moment and cannot be extended in time or, I think one should add, be repeated in a relatively short time-span.[9] As noted by, for example, Platzack (1979, 71), achievements constitute a special type of bounded event/act characterizable by the semantic unit [PUNCTUAL EVENT/ACT]. Events/acts characterized by the kind of finality under consideration, accordingly, are punctual events/acts too. But punctuality is an aspectual parameter, which means that the correspondence between punctuality and finality makes it possible to capture the essence of the generalization (35) by a conceptualization rule formulated in terms of aspectual notions. This rule may be formulated as (36).

(36) Conceptualize [ITERATED] in conjunction with the unit of aspectual meaning [NONPUNCTUAL] as [ITERATED AND ATTENUATED].

The conceptualization rule (36) says that the repetition of a nonpunctual event/act indicates its attenuation. [ATTENUATED], therefore, is a derived unit of meaning associated with reduplications whose verb bases have the aspectual meanings [BOUNDED] and [NONPUNCTUAL]. On the basis of (36), it is predicted that punctual verbs, i.e. verbs denoting achievements,

[9] For some discussion of Vendler's notion of "achievement" — and his related but distinct notion of "accomplishment" — cf. Lyons 1977, 711–712, Holisky 1981, 133ff., and Mourelatos 1981, 191ff. Accomplishments (such as to run a mile, to paint a picture, to grow up, etc.) and achievements differ in that the former, but not the latter, have intrinsic duration. Recently, Jackendoff (to appear, 7) has postulated the binary semantic feature [± Closed] to differentiate between Ordinary location (as expressed by *Some water was on the floor*) and Distributive location (as expressed by *Water was all over the floor*). He (p. 12), moreover, speculates that this feature "appears also in the semantic structure of Events, indicating this time *temporal* closure or lack thereof. Achievements and accomplishments (in the sense of Vendler 1967), which have a temporal endpoint, would be [+ Closed]; processes which are conceptualized without temporal endpoints, would be [− Closed]."

cannot be reduplicated in Afrikaans. This prediction is borne out by the semantic oddity of sentences such as (37a—d).

(37) (a) **Hulle* arriveer-arriveer *more.*
they arrive arrive tomorrow
'*They tentatively arrive a couple of times tomorrow.'
(b) **Tenzing* bereik-bereik *die kruin van Everest.*
Tenzing reach reach the summit of Everest
'*Tenzing tentatively reaches the summit of Everest a couple of times.'
(c) **Zola* wen-wen *die wedloop.*
Zola win win the race
'*Zola tentatively wins the race a couple of times.'
(d) **Hy* onthou-onthou *die voorval.*
he recall recall the incident
'*He tentatively recalls the incident a couple of times.'

Since the semantic unit [ATTENUATED] is derived by means of a conceptualization rule, it need not, and should not, be specified by the semantic interpretation rule (21).

5. Conclusion

Given the existence of devices such as the conceptualization rules that I have proposed, two conclusions may be drawn. The first is that the distinction between the meanings of "iteration", "continuation" and "attenuation" reduces to aspectual differences between verb bases of reduplications. The second and more general conclusion — which I have argued more fully elsewhere — is that as a formal means of word formation, reduplication in Afrikaans expresses one basic meaning only, viz. the meaning [INCREASED].[10] Thus, given the existence of conceptualization rules, only a single and strongly unifying semantic interpretation rule — the rule (21) — is required for all Afrikaans reduplications. Or, to put it the other way around, if we wish to understand the semantics of Afrikaans reduplication as a unitery phenomenon, we are driven to recognizing the existence of conceptualization rules. At this point various questions arise both from a language-independent and language-specific point of view about the nature of and further justification for conceptualization rules. I cannot deal with these questions within the

[10] I show elsewhere (Botha 1984, 149ff.) that there are compelling reasons to consider [INCREASED] rather than [DISTRIBUTED] as representing the basic meaning expressed by reduplication in Afrikaans.

restricted scope of this paper and for some discussion of them have to refer you to my more comprehensive study of Afrikaans reduplication.[11]

All I can do here is to point out that the conceptualization rules under consideration do not represent, from a general linguistic point of view, a unique kind of entity. Other linguists have independently proposed devices that are similar in content to conceptualization rules. I think in this connection of the principles that consitute the "logic of temporal relations" provided for by Miller and Johnson-Laird (1976) as well as by Lyons (1976). I think, moreover, of the rules of conceptual well-formedness postulated more recently by Jackendoff (1983). In (38) I present informal formulations by Miller and Johnson-Laird (1976, 444, 449) of two principles of the former sort, and in (39) similar formulations by Jackendoff (1983, 162) of two rules of the latter type.

(38) (a) "When you arrive somewhere, you stay for a while; when you reach somewhere you may or may not stay. So here is another difference in the temporal shape of verbs"

(b) "Thus BEGIN and END must entail R*t* [which is an operator that says merely that the state or process could be observed at some moment — R.P.B.], but they must say something more. A beginning has some sense that the event has not occurred before; an ending that it does not occur thereafter."[12]

(39) (a) ". . . the place function IN requires its reference object to be regarded as a bounded area or volume . . ."

(b) "The most salient place function expressed by 'on' requires its reference object to have an upper surface."

Finally, the conceptualization rules presented in the preceding sections may have the appearance of stipulations. A basic aim of further work has to be the reduction of these rules to more general principles of conceptualization.[13]

Bibliography

Andersson, S.-G.: What is aspectual about the perfect and the pluperfect in Swedish? In: De Groot and Tommola (eds), 199–207. 1984.

Botha, R. P.: On the "Galilean style" of linguistic inquiry. In: Lingua 58 (1982), 1–50.

Botha, R. P.: A Galilean analysis of Afrikaans reduplication. In: Stellenbosch Papers in Linguistics 13 (1984). Stellenbosch, Department of General Linguistics, U. S.

[11] For some discussion of these questions cf. Botha 1984, 151ff. and Botha 1986, 3.15.1.

[12] For a more formal reconstruction of some of the principles constituting the "logic of temporal relations" cf. Lyons 1977, 710.

[13] For further remarks on this possibility cf. Botha 1986, 153ff.

Botha, R. P.: Morphological form, semantic interpretation, conceptual structure: A Galilean analysis of Afrikaans reduplication. 1986. (Published as: Form and Meaning in Word Formation. A Study of Afrikaans Reduplication. Cambridge University Press, Cambridge 1988.
Bridgen, N.: Towards a functional grammar of aspect in Finnish. In: De Groot and Tommola (eds), 179—198. 1984.
Chomsky, N.: Aspects of the Theory of Syntax. MIT Press, Cambridge, Mass. 1965.
Comrie, B.: Aspect. An Introduction to the Study of Verbal Aspects and Related Problems. Cambridge University Press, Cambridge 1976.
Dahl, Ö.: On the definition of the telic-atelic (bounded-nonbounded) distinction. In: Tedeschi and Zaenen (eds), 79—90. 1981.
De Groot, C. and Tommola, H. (eds): Aspect Bound. A Voyage into the Realm of Germanic, Slavonic and Finno-Ugrian Aspectology. Foris Publications, Dordrecht 1984.
Heinämäki, O.: Aspect in Finnish. In: De Groot and Tommola (eds), 153—177. 1984.
Holisky, D. A.: Aspect theory and Georgian aspect. In: Tedeschi and Zaenen (eds), 127—144. 1981.
Holt, J.: Études d'aspect. In: Acta Jutlandica 15 (1943), 2.
Jackendoff, R. S.: Semantics and Cognition. Current Studies in Linguistics Series 8. MIT Press, Cambridge, Mass. 1983.
Jackendoff, R. S.: Distributive location. In: Sophia Linguistica (to appear).
Lyons, J.: Semantics 1—2. Cambridge University Press, Cambridge 1977.
Miller, G. A. and Johnson-Laird, Ph. N.: Language and Perception. The Belknap Press of Harvard University Press, Cambridge, Mass. 1976.
Mourelatos, A. P. D.: Events, processes, and states. In: Tedeschi and Zaenen (eds), 191—212. 1981.
Platzack, Ch.: The Semantic Interpretation of Aspect and *Aktionsarten*. A Study of Internal Time Reference in Swedish. Foris Publications, Dordrecht 1979.
Tedeschi, Ph. J. and Zaenen, A. (eds): Tense and Aspect. Syntax and Semantics 14. Academic Press, New York 1981.
Teleman, U.: Definita och indefinita attribut u nusvenskan. Studentlitteratur, Lund 1969.
Vendler, Z.: Linguistics in Philosophy. Cornell University Press, Ithaca 1967.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 39–52 (1986)

THE PRODUCTION AND INTERPRETATION OF *AD HOC* NOMINAL COMPOUNDS IN GERMAN: A REALISTIC APPROACH

HERBERT E. BREKLE

(University of Regensburg, FRG)

In 1977 on the occasion of the International Congress of Linguists in Vienna I proposed making an investigation of the conditions underlying the production, use and interpretation of *ad hoc* nominal compounds in German. In the meantime I and my colleagues have been able to carry out a large amount of resesearch in this area. The results have been pre-published on microfiche (approx. 2,500 pages, including working papers, experiments and corpora of data, about 3,000 *ad hoc*[1] nominal compounds found in written texts and about 400 from spoken texts).

One of the major premises of our approach was that a strict differentiation can and should be made between lexicalized and *ad hoc* nominal compounds; the former are assumed to be taken directly from the lexicon, the latter to be formed *ad hoc*, and in this respect to be comparable to sentences. This means, of course, that the task of the lexicon within a realistic grammatical model is a twofold one: to supply the lexical input for sentence formation directly from the list of morphologically simple and complex items *and* to generate morphologically complex structures on the basis of primary lexical material. The insertion of newly formed words in a sentence or a text depends on further rules and cannot in all cases be considered a strictly local process. The sheer wealth of data collected during the last six years indicates that in contemporary German — spoken and written — the formation of *ad hoc* nominal compounds is a highly creative and productive process. The majority of *ad hoc* nominal compounds are *not* formed in order to enrich the lexicon but are produced in order to secure the communicative flow of verbal exchange (depending on syntactic and semantic parameters to be specified). They thus have a rather ephemeral status comparable to that of sentences.

I cannot here go into details of the empirical side of our work (data-collecting, testing the *ad hoc*-ness of material, elicitation procedures for

[1] Published on microfiche as Brekle (1985a)

evaluative judgments in the semantic domain of nominal compounds and so on). Instead, I shall present some of our results.[2]

1. The structure of nominal compounds in German

Nominal compounds are complex words whose second member is a noun. According to the nature of the second member, or head, 5 possible types of *ad hoc* nominal compounds can be distinguished:

1. compounds with a synchronically non-derived head (e.g. *Pflanzentisch* 'plant table');
2. a lexicalised nominal compound as head (e.g. *Universitätswartezimmer* 'university waitingroom') or a denominal substantive (e.g. *Nachbar-Feindschaft* 'neighbour enmity');
3. compounds with a relational noun as head, that is, a non-derived noun with an argument-structure (e.g. *Auto-Bruder* 'car-brother');
4. compounds with deverbal heads (e.g. *Baumentscheidung* 'tree decision');
5. compounds with proper nouns as head (e.g. *Fenster-Schmidt* 'window Smith').

The first element of a compound can in principle contain words of any lexical category as well as (in the marked case) syntactic phrases such as NP, VP, Adv P, PP and S.[3] Coordinated structures of all types can also appear as the first member. Compounds, and in the present context especially *ad hoc* compounds, can be recursively embedded to give more complex, multiply *ad hoc* compounds. In the system of compounding we have taken as the basis for our study, all elements of the fixed lexical inventory (words and bound morphemes) carry information about their grammatical category as well as about the categories for which they subcategorize (if any). In addition, they contain information about their argument-structure, i.e., the syntactic category and semantic roles of the arguments they take. In the generative part of the lexicon, elements from this inventory are combined to form complex words, including compounds. The features and categories of these compounds are determined according to certain principles of the grammar, among which the

[2] The following sections of this paper are based on the final report ("Endbericht") of the project discribed in the opening paragraphs, namely Boase-Beier et al. (1985). Unless otherwise indicated, examples are taken from this report.

[3] Thus the restriction sometimes formulated to the effect that word-formation rules can involve no entities of a higher level than the word-level, cf. Selkirk (1982, 9), must be expressed in a weaker form: word-formations of this type constitute the marked case.

following is of central importance: the category and features of the right-hand member of a compound are assigned to the compound as a whole. This principle is based on the concept of "head" and the assumption (or Righthand Head Rule, cf. Williams (1981, 248), Selkirk (1982, 20)) that the head of a complex word in Germanic languages is the rightmost member. This restriction is to be seen as an extension of assumptions made in the Xbar-theory of syntax to the area of word-formation.[4] We regard formation-rules for words as being formally similar to phrase-structure rules of the syntax, though as being situated in the generative component of the lexicon.[5]

2. Processes in the interpretation of nominal compounds

The interpretation of nominal compounds involves among other things the assignment to the compound of the semantic relation whose arguments are the constituents of the compound. We should like to make a distinction between two types of interpretation processes; this distinction is to be viewed in terms of a theory of markedness.[6]

We assume that a construction is either locally interpretable, that is, interpretable with reference only to material it contains, or it is non-locally interpretable, that is, it can only be interpreted with the help of external material. Because we assume that the central core of the grammar is characterized by a tendency towards as little deviation as possible from the local interpretation, we would wish to correlate the two types of interpretation with corresponding values in a theory of markedness, whereby the local interpretation is seen as unmarked, and the non-local as marked.

The interpretation of a nominal compound AB consists in constructing a level at which AB is represented as F(A) (or F(B)) or as F(A, B). That is, if there is no relation present in the compound AB, one has to be "found" and assigned to the compound such that A and B; the constituents of the context-dependent compound, are arguments of this relation. This is the non-local interpretation. If, however, A or B is a relational element, it is necessary to decide in what relationship the other element stands to this relation. If the non-relational element is an argument of the relational one, the compound can be interpreted locally. The means used to construct the semantic representation vary according to the type of knowledge required. These types of knowledge in turn correspond to the types of information associated with the lexical entries for the two constituents of the compound.

[4] See Toman (1983).
[5] In this we follow Selkirk (1982).
[6] See Boase-Beier & Toman (1985).

Compounds can be divided into 4 classes according to the type of information associated with their constituents (and thus according to the means used to interpret them), namely: relational compounds, stereotype compounds, compounds with a basic relation and context-dependent compounds.

2.1. Relational compounds. These compounds contain a relational element (i.e. a verb or verbal derivation, a relational noun (such as *Bruder* 'brother'), a preposition or an adjective) as either the first or second member. In the lexicon, information is associated with this relational element about the arguments (i.e. their syntactic type and semantic roles) which it can take. In general, relational compounds are so interpreted that the non-relational element is an argument of the relational one, e.g. *Hausvermietung* 'renting of houses', or an adverbial modifier, e.g. *Wintervermietung* 'renting in winter', and are thus locally interpretable. Their interpretation requires only grammatical knowledge.

2.2. Stereotype compounds. Though these compounds do not contain an explicitly relational element, one of their members is implicitly relational in that a relation is associated with it in the lexicon. For example, in the case of a compound such as *Milchwagen* 'milk van', we assume that the relation 'transport' is given in the lexicon as a stereotype of the word *Wagen*. Other relations associated with words, such as "produce" in the case of *Fabrik* 'factory' are listed in the lexicon as part of the meaning of the word in question, rather than as a stereotype of it. Both types of relations inherently associated with nominal elements in compounds are here subsumed under the heading of "stereotype relations".[7] Compounds with one of whose members such a relation is associated may thus be interpreted in a way similar to that appropriate to relational compounds although their interpretation requires more than purely grammatical knowledge, because knowledge of stereotypic relations involves lexical-semantic information and also semantic-pragmatic knowledge of the world.

2.3. Compounds with a basic relation. These compounds are not locally interpretable; they can only be interpreted with recourse to information not contained in their constituents. The relation underlying a compound of this type is one of a small set of basic relations containing the relations "is the location for" (as in *Waldhaus* 'forest house'), and the relation of coordination (as in *Hausboot* 'house-boat'), and possibly the relation "is similar to" (as in *Butterblume* ('butter-flower') "buttercup"). Such relations cannot be obtained from the compound itself. The type of knowledge required to interpret these

[7] See Boase-Beier (1985a), Fanselow (1981).

compounds is very complex in nature, involving knowledge of the basic relations themselves in addition to grammatical knowledge about their argument structures and semantic knowledge about the stereotypic and other qualities associated with the constituents of the compounds.

2.4. Context-dependent compounds. Compounds which ar e not locally interpretable and are not assigned a basic relation are context-dependent; they can only be understood using information taken from the context in which they occur. Their interpretation presupposes knowledge about rules of text-construction which, together with the information the context contains, enable the listener to identify the correct relation in the context. In this group are three main types of compounds:

1. Non-relational nominal compounds without a stereotype-relation, e.g. the compound *Schlangenmann* 'snake man' which, as it has no inherent relationality, cannot be assigned any interpretation without recourse to context information;

2. relational nominal compounds which have a context-dependent reading, e.g. the compound *Autoverkäufer* 'car seller' meaning not a seller of cars but a seller (of something else) who comes to work by car;

3. stereotype compounds which have a context-dependent reading, i.e. *Bücherregal* 'book shelf' meaning not a shelf for books but a shelf (for something else) which is constructed of books;

In cases 2 and 3 the compounds cannot be interpreted on the basis of the argument-structure of their inherent relation, thus they, like 1, are context-dependent.

2.5. Empirical evidence. It is to be expected that these theoretical differences concerning the type of information connected with a compound will be manifested in various interpretation-strategies during the actual interpretation of *ad hoc* nominal compounds. Results of experiments which we have carried out show that compounds in these different groups are in fact understood at different speeds. The experiments appear to substantiate the view that compounds for which grammatical knowledge alone forms a sufficient basis for their interpretation are understood more clearly (and in terms of reaction-times, more quickly) than those whose interpretation requires additional semantic or pragmatic knowledge. A compound of the latter type will be understood more easily (and quickly) than one which can only be interpreted using additional context and text-grammatical knowledge.

3. Pragmatic aspects of nominal compounding

It is to be assumed that a complete and empirically adequate theory of nominal compounding must also contain reflections on its pragmatic aspects.[8] Taking pragmatics to be the study of the appropriateness of a particular utterance in a particular context, it is clear that the choice of a nominal compound in a particular context is influenced above all by considerations which can be summarized under the following two pragmatic aspects:

1. When is a nominal compound *suitable* to a particular context?
2. When is a nominal compound *understandable* in a particular context?

The first question concerns the production of *ad hoc* nominal compounds, the second their interpretation. The relation between interpretation and production can be given as follows: only interpretable compounds are produced. This is of course not to be seen as a statement specific to nominal compounding, but as a general principle for the use of linguistic structures. Thus all factors which allow nominal compounds to be interpreted are also directly relevant for their production, though the opposite clearly does not apply: factors which make the production of a nominal compound in a particular context likely need not necessarily contribute directly to its interpretation.

In what follows we shall briefly discuss pragmatic factors from both these aspects, i.e. in connection with both the production and the interpretation of *ad hoc* nominal compounds.

3.1. The production of nominal compounds. Apart from the question of individual style, there are a number of factors whose presence renders the production of nominal compounds appropriate. These factors are of varying status: some are context-features, some are special effects associated with *ad hoc* nominal compounds, some are specific aims which can be realized using nominal compounds. We shall consider some of these in the sections which follow.

3.1.1. Lexical gaps. There are two main reasons for the existence of a lexical gap: either there is objectively a gap in the lexicon of the language, or the speaker himself does not have access (either in general, or as the result of a momentary weakness in performance) to the required word.[9]

In general, every use of an *ad hoc* compound (or of any other new word formation) presupposes a lexical gap of some sort, in the sense that the for-

[8] Cf. Brekle (1978, 76) and also Herbermann (1979), Mabrey (1980).

[9] Cf. the "Semantic Gap Hypothesis" put forward by Carroll and Tanenhaus (1975, 51).

mation of a new word is only possible if no word already exists in the lexicon with exactly the same meaning.[10] Thus the formation of *ad hoc* compounds which at first sight appear to be synonymous with lexicalized words can be seen as the consequence of a "stylistic gap". Nominal compounds may thus serve the potential extension of the lexicon in the sense that a phenomenon which had not existed or not been apprehended before is given a name. Examples of this are the compounds *Holocaust-Film* 'holocaust-film', *Tschernobyl-Wolke* 'Chernobyl-cloud' and *Makrobioten-Restaurant* 'macrobiotic restaurant'. But gaps also exist in the system of a language; some lexicalized expressions, for example, only occur in the plural. The formation of a new compound can then render it possible to use the expression in the singular form. An example is the expression *die Flammen der Hölle* 'the flames of hell', corresponding to which the *ad hoc* compound (eine) *Höllenflamme* 'a hell flame' can be formed.

In this connection the phenomenon of reification should also be mentioned. The reifying effect of *ad hoc* nominal compounds is used in scientific terminology, and also frequently in the service of ideologies.[11] An example of the reifying effect of such compounds can be seen in the case of the compound *Batterie-WAA* (*Wiederaufarbeitungsanlage*) 'battery reprocessing plant', which was formed in the context of a talk on atomic energy reprocessing plants and environmental protection. There exist no battery reprocessing plants, but the idea is put forward as worthy of consideration. The use of the compound serves to reify the concept of such an enterprise.

The results of an experiment carried out by us indicate the relevance of the concepts of the lexical gap for the formation of *ad hoc*-compounds. The results further showed that the formation of compounds, given the existence of lexical gaps, is influenced by two factors — the familiarity and the complexity of the object to be named. Unfamiliar and complex objects led to the formation of more compounds than did familiar and non-complex objects.

3.1.2. Reference. Under this heading can be subsumed both the quasi-pronominal use of nominal compounds and the "deictic" use.[12] The quasi-pronominal use is considered by various linguists to be the main function of word-formation.[13] Compounds often occur in a text when the use of a pronoun would not suffice to establish referential identity between two expressions. The following example illustrates this:

[10] Cf. also the principle of Pre-emption, as put forward by Clark and Clark (1979).
[11] See Brekle (1985b).
[12] Cf. Downing (1977, 339).
[13] Cf., for example, Brekle (1985b, 18f.).

(1) *die Wahlkampfmannschaft von Strauß . . .*
. . . die S t r a u ß - M a n n s c h a f t[14]
'Strauß's election-campaign team . . .'

By "deictic" reference is meant the use of an *ad hoc* compound which makes it possible for the speaker to refer to something in a way which is clear and comprehensible. This use, described in detail by Downing (1977) was tested by us in an experiment and it was found that in the referential use the proportion of *ad hoc* nominal compounds was very high.

3.1.3. Contrast. Contrast in the context of an *ad hoc* nominal compound plays an important role in its interpretability, and is thus also a significant factor in the production of such compounds. As the latter aspect will be considered later, we shall here only mention the former. A nominal compound can be formed in contrast to a lexicalized compound or phrase, or to an expression which precedes it in the context. For example, in contrast to the lexicalized compound 'family allowance', the compound *Elterngeld* 'parents' allowance' can be formed. In the following example, a compound is formed in contrast to one formed earlier in the text:

(2) *Im Münchner Zoo lebt ein Adler, der nur Rosen frißt. Der R o s e n a d l e r ist jedoch unter den Besuchern nicht so beliebt wie ein südafrikanischer T u l p e n a d l e r.*
'In the Munich Zoo there is an eagle which only eats roses. However, the r o s e e a g l e is not so popular among visitors as a South African t u l i p e a g l e.'

Both these compounds are *ad hoc*. The first is explained in context and the second is understood as a contrast to the first. Even compounds which would normally not be considered acceptable tend to be accepted when there is a clear contrast present in the cotext.

3.1.4. The Minimax Principle. A further condition for the use of *ad hoc* nominal compounds can be seen in the so-called Minimax Principle,[15] which says that a speaker will always try to form utterances such that minimum surface complexity is combined with maximum information content. Compounds are suitable means of fulfilling this general communication principle. If we use, for example, instead of *das Problem des Straßenbaus* 'the problem of building streets', the compound *Straßenproblem* 'street-(building) problem',

[14] Example taken from Wildgen (1985).
[15] Compare Carroll and Tanenhaus (1975, 51).

we are acting according to this principle. The results of an experiment, in which argument structures present in a text were to be used again later in the same text showed that the structures most commonly used were compounds, suggesting that the Minimax Principle influences the choice of construction-type.

3.1.5. Foregrounding. This is a term originally used by the Prague School Linguists[16] and we mean by it the directing of the attention of the listener or, more often, the reader, to the form of an utterance, rather than to its content. All *ad hoc* words, by virtue of being innovative in nature, have this effect to some extent. The effect is used above all in poetic language, in advertising, and in journalism, particularly in headlines. It is important for the production of *ad hoc* nominal compounds in so far as the aim of drawing attention to a linguistic form often results in the use of *ad hoc* compounds.

3.1.6. Ambiguity. All *ad hoc* compounds without an explicitly expressed relation are ambiguous, because in their semantic representation in principle any relation could be assigned to them. Thus a compound *Blatt-Schmetterling* 'leaf-butterfly' could mean 'a butterfly which looks like a leaf', 'a butterfly which lives on leaves' and so on. Normally such a compound is disambiguated in context. However, the inherent ambiguity of such forms can be used for particular effects and thus the intention of achieving ambiguity is a condition which makes the use of a compound more likely. This potential ambiguity is used particularly in political contexts but also in poetic language.

3.1.7. Analogy. The importance of analogy in the formation of new words has been observed by many linguists.[17] A lexicalized compound serves as a pattern for a further compound or series of compounds. Thus, for example, parallel to *Augenarzt* 'optician', *Zahnarzt* 'dentist', we can form the compound *Beinarzt* 'leg-doctor' for a doctor who treats the leg of his patient. Not only lexicalized compounds, but also *ad hoc* compounds can serve as the basis for analogical new forms within the same text.

The role of analogy in the production of *ad hoc* nominal compounds has been empirically tested and it was found that the presence of a lexicalized or new compound in a text resulted in a greatly increased production of analogous *ad hoc* compounds.[18]

3.2. The interpretation of nominal compounds. The production of *ad hoc* nominal compounds is influenced by all the above factors, which often appear

[16] See for example, Mukařovský (1932 [1964]).
[17] See, for example, Paul (1896), Motsch (1977), Fleischer (1969).
[18] See Stöhr (1985).

in combination with another. But the interpretability of *ad hoc* compounds must also be ensured.

For the interpretation of locally non-interpretable structures such as non-relational compounds, we can formulate the following Principle of Interpretation:[19]

(3) Given a locally non-interpretable construction, attempt to apply the interpretational mechanisms provided by the grammar such that a semantic representation can be assigned to the construction.

Corresponding to this, a Pragmatic Principle could also be formulated with the form: "form only such constructions as are interpretable".

As mentioned above, the interpretation of every compound which is not locally interpretable involves finding a relation which can be assigned to the compound AB to give a semantic representation R(A, B). In order for this relation to be found in the context, there must be present two elements, A' and B', which correspond to A and B, and which are connected by a relation. The relation obtaining between A' and B' in the context is then the relation to be assigned to AB at the semantic level. A principle such as (3), ensuring that the relation R is assigned to the compound, would of course interact with a number of other principles which ensure that the correct relation can be assigned. In this connection, Grice's conversational maxims are relevant[20] Grice mentions four main maxims, namely the maxims of Quantity ("say as much as is necessary"), Quality ("speak the truth"), Relation ("be relevant") and Manner ("be clear"). For the interpretation of an *ad hoc* nominal compound in context, these maxims mean respectively, among other things:

a) the relation to be assigned to the compound must be present in the context;
b) the relation to be assigned to the compound must be identifiable in the context;
c) the compound must be relevant in the context;
d) the compound must be applicable (we cannot speak of truth in connection with a lexical item).

3.2.1. The Presence of the Relation. If a compound *Grammatik-Kleid* 'grammar dress' is to be interpretable, both the relation and two arguments which correspond to the elements of the compound must be present in the context:

[19] See Boase-Beier and Toman (1985), here translated.
[20] See Grice (1975).

(4) *Alle Mädchen des Wolfgangs-Gymnasiums besitzen die englische Grammatik mit dem leuchtend grünen Einband. Als die Englisch-, Lehrerin ein auffallendes Seidenkleid in einem ähnliche Farbton trug, sagten die Mädchen, sie hätte sich ein G r a m m a t i k - K l e i d gekauft.*
'All the girls of the Wolfgang Grammar School own the English Grammar with the bright green cover. When the English teacher wore a similar colour, the girls said she had bought a g r a m m a r d r e s s.'

In this example, as much information as is required to interpret the compound, in the form of the two corresponding elements A' (*Grammatik*) and B' (*Kleid*), as well as the relation "like" (i.e. *in einem ähnlichen Farbton*) is present in the context. If this information is not present, the compound is not interpretable:

(5) *Frau Schmidt trug ein Grammatik-Kleid.*
'Mrs Smith wore a grammar dress.'

Here there are no elements A', B' corresponding to the elements of the compound and thus no relation can be assigned.

3.2.2. Identifiability In the above example (4), the relation is identifiable because it clearly appears in the context in the argument structure "like (dress, grammar)".If the relation is not identifiable, the compound cannot be interpreted; compare the following example:

(6) *Am Montag, an dem Tag, an dem in unserem Laden immer sehr viel verkauft wird, sprach ich mit der K a r t e n f r a u.*
'On Monday, the day on which a great deal is always sold in our shop, I spoke to the card woman.'

Suppose that the compound *Kartenfrau* 'card woman' is to receive the reading 'woman who sells cards'; the relation *verkaufen* 'sell' appears in the context but there is no way of identifying it as the relation between *Karten* and *Frau*. Assuming that the finding of the correct relation in the context involves not only the presence of the corresponding argument structure in the context, but also the means to identify it, we can observe a number of restrictions on the relation between the corresponding elements A' B' in the text and the elements of the compound themselves. These restrictions are discussed in detail in Boase-Beier and Toman (1985), and will just be mentioned briefly here: they are the "No-Mixing Principle" and the "Specific-General Order Prin-

ciple''. The former says that the elements of the compound must be arguments of one and the same relation, e.g. a compound *Rosenwald* 'rose forest' may not be interpreted as 'forest in which there lives a magician who eats roses'. The unacceptability of compounds with such readings was substantiated in empirical tests. The Specific-General Order Principle states that, for the interpretation of *ad hoc* nominal compounds, a compound with quasi-pronominal reference can only be understood if the order ''specific-general'' is observed, i.e. if the elements of the preceding text which correspond to the elements of the compound are more specific than the elements of the compound themselves.

3.2.3. Relevance A compound which is not relevant in context cannot be interpreted. An *ad hoc* compound *Glasmuseum* 'glass museum' is locally interpretable as 'museum in which glass is displayed'. But in the following text, the compound appears not to be relevant:

(7) *Als Hans neue Zimmerpflanzen kaufen wollte, besuchte er zuerst das G l a s m u s e u m und fuhr dann in die Stadt zu einer Gärtnerei.*
'When Hans wanted to buy new plants for his room, he first visited the g l a s s m u s e u m and then went to town to a garden centre.'

The listener assumes the speaker will say what is relevant. Thus he will assume the normal reading of *Glasmuseum* does not apply here and will try to interpret it so that it has something to do with plants.

In empirical tests it was found that compounds were in general interpreted as relevant to their contexts. The same compound, even if locally (i.e. context-independently) interpretable, was interpreted diffently in different contexts, so that it was in each case relevant to its context.[21]

3.2.4. Applicability. In the following example, the compound is relevant in the context, which contains sufficient information to interpret it. However the compound appears not to be applicable.

(8) *Ein Laden, in dem Pflanzen verkauft werden, heißt F i s c h l a d e n.*
'A shop in which plants are sold is called a fish shop.'

In a context like this, the compound remains uninterpretable.

3.2.5. Pre-emption. The principle of Pre-emption, as formulated by Clark and Clark (1979), says that a new lexeme can only be formed when there is no lexeme with the same meaning already present in the lexicon. Thus in general

[21] See Beier (1985).

a new compound *Grasmäher* 'grass mower' will not be formed to denote that which is normally called *Rasenmäher* 'lawn mower'. From the point of view of interpretation this means that, when a new compound is formed it will be interpreted to mean something different from the already existing word, unless the listener in an actual conversation assumes the speaker is suffering from a subjective lexical gap where objectively there is none. If, however, he has reason to assume that this is not the case, he will interpret the compound as having a new meaning; that is, either as denoting a new object — in the case above perhaps a special mower for high grass — or as having a different stylistic or connotative value from that of the lexicalized compound.

3.2.6. Analogy. We have spoken above about the role of analogy in the production of compounds. It is of course also of importance in their interpretation. Especially for compounds which are not locally interpretable, such as compounds whose underlying relation belongs to the set of basic relations, analogy plays an important role. For example, the compound *Blutpflaume* 'blood plum' could be interpreted as a plum with reddish flesh in analogy to the lexicalized compound *Blutorange* 'blood orange'. As the above mentioned experiment on the role of analogy in the production of *ad hoc* nominal compounds showed, analogy to a lexicalised compound was the strongest single factor influencing the interpretation of an *ad hoc* compound.[22]

From the above discussion it will be clear that a thorough study of the production and interpretation of *ad hoc* nominal compounds involves the investigation of a large number of factors, syntactic, semantic and pragmatic. It is to be hoped that this paper has served to provide some insight into the type of questions involved.

References

Beier, D.: Experimentelle Untersuchung zur Interpretation eines ad hoc-Nominalkompositums in verschiedenen Kontexten. In: Brekle, H. E. (ed.) 1985a.

Boase-Beier, J.: Die Semantik von ad hoc-Komposita mit zwei nominalen Bestandteilen. In: Brekle, H. E. (ed.) 1985a.

Boase-Beier, J. et al.: Endbericht des DFG-Projekts Nominalkomposita. In: Brekle, H. E. (ed.) 1985a.

Boase-Beier, J. and Toman, J.: Komposita und Kontext. In: Brekle, H. E. (ed.) 1985a.

Brekle, H. E.: Reflections on the conditions for the coining, use and understanding of nominal compounds. In: Dressler, W. U. and Meid, W. (eds): Proceedings of the 12th International Congress of Linguists. Innsbruck 1978, 68—77.

Brekle, H. E. (ed.): Kommunikative und pragmatisch-semantische Bedingungen der Aktualgenese, der Verwendung und des Verstehens von Nominalkomposita im Deutschen. Teil 2. Regensburg Mikrofiche Materialien No. 012, Regensburg 1985a.

Brekle, H. E.: Bedingungen für die Aktualgenese deutscher Komposita. In: Brekle, H. E. (ed.) 1985a.

[22] See Stöhr (1985).

Brekle, H. E. and Kastovsky, D. (eds): Perspektiven der Wortbildungsforschung. Beiträge zum Wuppertaler Wortbildungskolloquium vom 9—10. Juli 1976. Bouvier, Bonn 1977.

Carroll, J. M. and Tanenhaus, M. K.: Prolegoma to a functional theory of word formation In: Grossman, R. E. et al. (eds): Papers from the Parasession on the Lexicon. Chicago Linguistic Society, Chicago 1975, 47—62.

Clark, H. H. and Clark, E. V.: When nouns surface as verbs. In: Language 55 (1979), 767—811.

Downing, P.: On the Creation and Use of English Nominal Compounds. In: Language 53 (1977), 810—842.

Fanselow, G.: Zur Syntax und Semantik der Nominalkomposition. Ein Versuch praktischer Anwendung der Montague-Grammatik auf die Wortbildung des Deutschen. Linguistische Arbeiten 107. Niemeyer, Tübingen 1981.

Fleischer, W.: Wortbildung der deutschen Gegenwartssprache. VEB Bibliographisches Intitut, Leipzig 1969.

Grice, H. P.: Logic and Conversation. In: Cole, P. and Morgan, J. (eds): Speech Acts. Syntax and Semantics 3. Academic Press, New York 1975, 41—58.

Herbermann, C. F.: Wort, Basis, Lexem und ihre Rolle bei der Bildung komplexer Substantive. Habilitationsschrift, Bochum 1979.

Mabrey, C.: Pragmatic-Semantic Constraints for the Coining of ad hoc Nominal Compounds in German. MA-Thesis, University of Regensburg 1980.

Motsch, W.: Ein Plädoyer für die Beschreibung von Wortbildungen auf der Grundlage des Lexikons. In: Brekle, H. E. and Kastovsky, D. (eds). 1977, 180—202.

Mukařovský, J.: Standard Language and Poetic Language. In: Garvin, P. (ed.), 1964 17—30. (First published in Czech 1932).

Paul, H.: Über die Aufgaben der Wortbildungslehre. Sitzungsberichte der philosophisch-philologischen und der historischen Classe der königlich bayerischen Akademie der Wissenschaften zu München 1896, 692—713.

Selkirk, E. O.: The Syntax of Words. Linguistic Inquiry Monographs, 7. MIT Press, Cambridge, Mass. 1982.

Stöhr, I.: Experimentelle Untersuchung zur Rolle kotextueller Faktoren bei der Produktion von ad hoc Nominalkomposita. In: Brekle, H. E. (ed.), 1985a.

Toman, J.: Wortsyntax: Eine Diskussion ausgewählter Probleme deutscher Wortbildung. Linguistische Arbeiten 137. Niemeyer, Tübingen 1983.

Wildgen, W.: Kompositionsprozesse am Beispiel ad hoc gebildeter Nominalkomposita im Deutschen. In: Brekle, H. E. (ed.), 1985a.

Williams, E.: On the Notions "Lexically Related" and "Head of a Word". In: Linguistic Inquiry 12 (1981), 245—247.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 53–67 (1986)

HOW TO FIX INTERFIXES? ON THE STRUCTURE AND PRAGMATICS OF ITALIAN (AND SPANISH, RUSSIAN, POLISH) ANTESUFFIXAL INTERFIXES AND OF ENGLISH "INTERMORPHIC ELEMENTS"

WOLFGANG U. DRESSLER, (Vienna)

LAVINIA MERLINI BARBARESI (Parma)

Introduction

1. Dressler (1984; 1985b, 329—330; 1986) dealt with the universal unnaturalness, typological adequacy and language specific aspects of interradical interfixes such as E. *chem-o-therapy*, Span. *alt-i-plano*, G. *Mensch-en-affe* and of mainly Spanish and Russian antesuffixal interfixes such as Sp. *casa* 'house' → augmentative *cas-er-ón*, *tos* 'cough' → adj. *tos-eg-oso* Russ. *Glinka* → adj. *Glink-ov-skij*, *Glink-in-skij*. There he also believed Tekavčić's (1968) results that Italian had no interfixes. This generalization seemed unacceptable to Merlini so that she proposed to write a coauthored paper on the subject. Simultaneously Szymanek (1985a, b) put intermorphic elements such as E. *-at-*, *-u-* in *dram-at-ical*, *process-u-al* and Polish interfixes on a par. Both authours of this paper immediately disagreed with this view.

2. From our consideration of interfixes we must first exclude a) infixes such as *-n-* in Lat. *fra-n-g+o* vs. *frac+tus*, *frag+mentum*, b) inflectional class indices such as in Lat. *flor-e-o* vs. incohative *flor-e-sc-o*, c) isolated insertions such as in the adjective of *Rousseau* Sp. = It. *rousseau-n-iano*, d) predictable insertions by an ordinary allomorphic or morphonological rule (see Dressler 1985a), be it within the word domain as of *-t-* in Fr. *abri* → *abri-t-er* or within the phrase/sentence domain as in Fr. *a-t-il*, *étai*[t]*-il*, Fr. avancé: *j'étai*[t] *ici*.

Descriptive Comparison of Spanish,[1] Russian,[2] Polish,[3] Italian[4] and English[5]

3. First let us compare properties of (briefly) Spanish, Russian, Polish and (more extensively) Italian antesuffixal interfixes on the one hand with Szymanek's (1985a, b) putative English intermorphic elements on the other.

3.1. First criterion: All Spanish, Russian and Polish interfixes are meaningless, i.e. semantically empty. The same criterion holds for English putative interfixes. Italian candidates, interfixes *-ic-*, *-ar-/-er-*, *-ol-* before diminutive suffixes, do not have a denotative meaning, but they may contribute a connotative meaning (cf. see below **5**, 11). For a few more marginal antesuffixal interfixes see Prati (1942).

3.2. Second criterion: One and the same interfix recurs after many different bases and before many different suffixes, and various interfixes recur after the same base and before the same suffix, e.g.

(1) Sp. *polvo* → *polv-ar-eda*, *humo* → *hum-ar-eda* vs. *pera* → *per-eda*, *tufo* → *tuf-ar-ada*, *bulla* → *bull-ar-anga*, *hoja* → *hoj-ar-asca*, *monte* → *mont-ar-az*, *espuma* → *espum-(ar)-ajo*, *vivo* → *viv-ar-acho*, *dicho* → *dich-ar-acho*.

Due to this free recurrence of interfixes it would be, at least, very uneconomic to assume many similar variants of bases and of suffixes respectively instead of a small number of identical interfixes. This is a classic case of morphological segmentation and identification despite lacking semantic criteria for identification (cf. Zemskaja 1964, Aronoff 1976). The same holds for Italian candidates as in:

(2) It. *volta* → *volt-ic-ina*, *volt-ar-ella*, *volt-ic-ella*; *vetta* → *vett-ar-ella*, *vett-ic-ina*, *vett-ol-ino*; *banco* 'bench' → *banc-ar-ello*, *banch+er-ott-olo*, *banch-er-ozzo*; *prato* → *prat-ic/er-ello*, *prat-ic-ino*; *pianta* → *piant-ic-ina/ella* (all diminutives); *carne* 'meat' → adj. *carn-ic-ino*; *paglia* 'straw' → *pagli-er-ino*, *marzo* 'march' → *marz-ol-ino*; *nodo* 'knot' → *nod-(er)-oso*; *freddo* 'cold' → *fredd-ol-oso*.

[1] After Malkiel 1958 etc. and Dressler (1986) and Dressler's work with 3 informants.
[2] After Zemskaja 1964, Lopatin 1975, Dougherty 1984 etc. and Dressler's work with 3 informants.
[3] After Szymanek 1985a, b and Dressler's work with 3 informants.
[4] After Merlini's research, cf. Prati 1942. More at the Bologna meeting of the Società di Linguistica Italiana (1986).
[5] After Merlini's research, more in a forthcoming article of hers.

As to E., let us consider Szymanek's (1985a, b) putative interfixes *-l-*, *-v-*, *-n-*, *-at-*, *-i-*, *-u-*, *-m-*; *-l-* and *-v-* are totally isolated stem variants of *Congo* and *Peru*[6] in *Congo-l+ese* and *Peru-v+ian*, and can be accomodated with allomorphy rules. The same applies to *-n-* after classical names and toponyms such as, *Cicero-n+ian* > *Toronto/Buffalo-n+ian, Panama-n+ian, Java-n+ese.* An allomorphy rule (in the sense of Aronoff 1976, or an allomorphic rule in the sense of Dressler 1985a) applies also to Greek words in *-ma* and *-asis* before the suffix *-ic* as in *drama, schema, idiom* → *dram-at-+ic, schem/idiom-at+ic, psoriasis* → *psori-at+ic, emphasis* → *emph-at+ic.* Also note *derm+al, derm+ic* vs. *derm-at+itis, derm-at+o+logy, derm-at+o+gen* with the interradical interfix *-o-* which never follows an antesuffixal interfix. Of course *derm-ic,* does not have the enlargement *-at-* because there is no autonomous English word **derm.* As to

(3) *term-in+al, term-in+ate, co(n)term-in+ous; pag-in+al/ate; crim--in+al, incrim-in+ate; attitud-in+al/ize, latitud-in+al/ar-ian, multitud-in+ous, longitud-in+al*

an allomorphy rule adds *-in-* to a few words such as *term, page, crime, germ* and to all words in *-tude. Crim-in+o+logy* shows again that *-in-* is no interfix because it is followed by the interfix *-o-*. Better candidates for iterfixal status are *-u-* and *-i-*; *-u-* as in

(4) *percept+u-al, habit/act/accent+u-al/ate; sens+u-ous, contempt/volupt/(pre)sumpt/tumult+u-ous/ary, text+u-al/ary.*

It is added to latinate bases before the suffixes *-l, -ous, -ate, -ary.* The only non-latinate base occurs in *theft+u-ous. -i-* appears frequently after latinate bases before suffixes *-al, -ous, -an, -ate*:

(5) *baron/manage/agent/part+i-al; labor+i-ous, sentent/pretent+i-ous, (obv+i-ous); Corinth+i-an, microb/magic/civil/Darwin/Paris+i-an; professor-i-ate, margrav+i-ate, vicar+i-ate.*

The variant *-i-an* is obligatory after *-l,* as in *mammal+ian, camel+ian, reptil+ian* and the default after *-ic* as in *phonet/statistic/ballistic+ian,* the only 4 counterexamples to be found in Lehnert (1971) are *(re)public+an, suburbic+an, Gallic+an*; *-ial* is the obligatory variant of *-al* after orthographic ⟨ce⟩ as in

[6] *Shaw — Shav-ian* is a case of weak suppletion whereas *Craco-v+ian* shows like *Peru-v+ian* an intrusive *-v-* after the base (*Cracow*).

(6) *spac+ial, fac+ial, artific+ial, edific+ial*
and is the default after a suffix of the precise form *-ent*:

(7) *presid+ent+ial, exist+ent+ial,* but *anteced+ent+al, coincid+ent +al, transcend+ent+al, resid+ent+(i)al* vs. *dent+al, percent+al, parent+al, client+al, depart+ment+al, environ/incre/sedi/govern+ ment+al.*

-ial is also the default after *-or/-er* with a few variants *-(i)al*, cf. the following bases:

(8) *-ial*: *sponsor-, professor-, ambassador-, examinator-, investigator-, minister-(i)al*: *rector(y?)-, inspector-, tector-, doctor-, elector-, prefector-, soror-, sector-. -al*: *flor-, chlor-, femor-, humor-, tumor-, tempor-, spor-, corpor-, peripher-, later-, armiger-, carnivor-.*

Thus the recurrence of the putative interfixes *-i, -u* is less free than in Italian, Spanish, Russian, and Polish. Therefore we propose to handle them in terms of suffix variation.

3.3. Third criterion: Spanish, Russian, Polish and Italian interfixes have such a degree of productivity that we can assume productive or at least semiproductive word formation rules.
They occur with intralingual neologisms such as in

(9) It. *scontro* 'car crash' → *scontr-ic-ino, busta* 'envelope' → *bust-ar-ella* 'bribe', *cocco* 'darling' → dim. *cocch-ino* vs. *Cocc-ol-ino* 'brand name of a laundry softener' or with recent loan-words, e.g.

(10) Pol. *kwakier* 'quaker' → adj. *kwakier-ski* = *kwakr-ow-ski, szvagier* 'brother-in-law' → *szvagier-ski* = *szvagr-ow-ski, klown* → *klown-ow- -ski, bonza* → *bonz-ow-ski*

and with abbreviations, e.g.

(11) Pol. *RFN* 'West Germany' → *erefen-ow-ski*; Russ. *TASS* → *tass-ov- -skij, VUZ* → *vuz-ov-ec, MGU* → *èmgeu-š-nik. ASEAN* → *asean-ov- -skij*; Sp. *ONU* → *onu-s-iano, PCE* → *pece-c-ito, PSOE* → *psoe-t-azo, CIA* → *ciá-t-ico.*

In English the suffix variants *-ial* and *-ian* are productive with Latinate bases and with names, as in *manager-ial, entrepreneur-ial, statistic-ian, Keynes-ian,*

Nixon-ian. Neither in Italian nor in English there are derivations from abbreviations.

3.4. Fourth criterion: The insertion of a Spanish, Russian or Polish interfix is only partially predictable, i.e. with some probability, e.g. in Spanish the most productive interfix is *-ec-* before diminutives, e.g. *pueblo* 'village' → *puebl-ec-ito* or *puebl-ec-illo*, but there is also *puebl-ito*, without interfix, but apparently no **puebl-illo*, however the situation is not identical with other bases before the same suffixes.

Italian interfixes are also unpredictable with the only exception of stems in *-on-* before the dimunitive suffixes *-ino*, *-ello* where a *-c-* must be inserted:

(12) *leone* 'lion' → *leon-c-ino*, *porta* 'door' → *port-one* → *port-on-c-ino*, *garzone* 'pot-boy (mixer)' → *garzon-c-ello*.

As to English we have seen that the choice of stem and suffix variants is nearly always predictable.

3.5. Fifth criterion: One condition favoring the insertion of an interfix in Spanish, Russian, and Polish (cf. Jaeggli 1980, Dougherty 1984, Szymanek 1985b) is a base/input condition: monosyllabic (and simultaneously monomorphemic) stems prefer the adjunction of an interfix:

(13) Sp. *madre* → *madr-ec-ita* vs. *co-madre* → *co-madr-ita*, cf. (*com*)*padre*, (*entre*)*mes*; *tos* 'cough' → adj. *tos-eg-oso*, dim. *tos-ec-illa*.

R. *slog* 'syllable' → adj. *slog-ov-oj* vs. *odno-složnyj* 'monosyllabic'
most 'bridge' → adj. *most-ov-oj* vs. *pred-most-nyj* 'before the bridge'
xod 'way' → adj. *xod-ov-oj* vs. *ras-xod-nyj* 'debtor's-'

Pol. *lord* → *lord-ow-ski*, *klown-ow-ski*; *stok* 'slope' → *stok-ow-y* vs. *Biał-o-stoc-ki*; exceptions: *pán-ski*, *car-ski*.

Suffixation to Italian monosyllabic (and simultaneously monomorphemic) words does not trigger interfixations, see

(14) *re* 'king' → dim. *re-uccio*, *tè* 'tea' → *te-ino*, cf. *gnu-ino*, *gru-ina*, *sci-ini*

But trisyllabic words (with two counterexamples) avoid the insertion of interfixes:

(15) *serpe* = *serpente* 'serpent' → dim. *serp-ic-ina/ella* vs. *serpent-ello*; (*rias*)*sunto* 'summary' → *sunt-er-ello* vs. *riassunt-ino*, but: *qual # cosa* 'something' → *qualcos-er-ella*, *volontà* 'want' → adj. *volont-er-oso*.

This condition does not obtain in English: There is even a counterexample: *verb-al* but *proverb-ial*, *adverb-ial*.

3.6. Sixth criterion: As consequence of criteria 1) and 4) there are synonymous doublets with and without interfixes,

(16) Pol. *kwakier-ski* = *kwakr-ow-ski*; R. *plug i molot-(ov)-skij*, 'plow-and-hammer-y', *mužik* 'peasant' → adj. *mužik-ov-skij*; Sp. *dal-ita* = *dali-ec-ita*, *yerb-(ec)-ita*, *gorr-(et)-ada*, *terr-(egu)-ero*, *negr-(eg)-or*, *negr-(eg)-ura*, *tenebr-(eg)-ura*; It. *topo* 'mouse' → dim. *top-(ol)-ino*, *corda* 'rope' → *cord-(ic/ol)-ina*, *nervo* 'nerve' → *nerv-(ic/ol)-ino*, *pianta* 'plant' → *piant-(ic)-ina*, *porta* → *port-(ic)-ina*

As to English, we have already seen the scarcity of variants in (8).

3.7. Seventh criterion: On the other hand there exist doublets with and without interfix where there is an (unpredictable) lexicalized meaning difference:

(17) Sp. *agua* → *agu-an-oso* 'watery' vs. *agu-oso*, *mano* → *man-ot-ada* 'hit with the hand' vs. *man-ada* 'flock',

R. *dom* 'house' → augm. *dom-ina* vs. *dom-ov-ina* 'part of the house'; *krug* 'cercle' → adj. *krug-ov-oj* vs. *kruglyj* 'round'; *krest* 'cross' → dim. *krest-ik* vs. *krest-ov-ik* 'garden spider'; *nos* 'nose' → dim. *nos-ik* vs. *nos-ov-ik* 'laryngologist'.

In Russian and Polish, names of vodka can be derived from names of berries and fruits with R. *-ov-ka* = Pol. *-ów-ka*:

(18) R. = Pol. *malina* 'raspberry' → dim. *malin-ka* vs. R. *malin-ov-ka* = Pol. *malin-ów-ka* 'raspberry vodka'; Pol. *orzech* 'walnut' → dim. *orzesz-ek* vs. *orzech-ów-ka* 'walnut vodka'.

Clearly — despite of the existence of a series of parallel derivations — the interfix does not mean 'vodka'. In the previous cases (17) there are even no semantically parallel derivations. The same holds for Italian:

(19) *corpo* 'body' → dim. *corp-ic-ino* vs. *corp-ino* 'bodice'
fante 'foot soldier' *fant-ic-ino* *fant-ino* 'jockey'
carta 'paper' *cart-ina* *cart-ic-ino* 'list of errata'
cart-ol-ina 'post card'

In English, however, the only conceivable examples would be *rector-al* (of God) vs. *rector-ial*, *sector-al* vs. *sector-ial* ('related to 'a line' vs. 'having the shape of a sector; related to a line').

3.8. Eighth criterion: Spanish and Russian interfixes may distinguish otherwise homonymous derivations, i.e. they differentiate outputs, e.g.

(20) Sp. *nueve* '9' → dim. *nuev-ito* vs. *nuev-ec-ito* ← *nuevo* 'new'; *corto* 'small' → *cort-ito* vs. *cort-ec-ito* ← *corte* 'court'; R. *det-* 'child' → adj. *det-skij* vs. *ded-ov-skij* ← *ded* 'grandfather'; *noga* 'foot' → *nož-noj* vs. *nož-ov-oj* ← *nož* 'knife'; *zenit* → *zenit-nyj* vs. business *Zenit* → *zenit-ov-ec*; *orel* 'eagle' → *orl-in-yj* vs. town *Orel* → *orl-ov-skij*; Pol. *krosno* 'spinning wheel' → adj. *krosienn-y* vs. place-name *Krosno* → *krosień-ski* = *krosn-ien-ski*

Italian cases are of several types: (21) gives examples of different interfixes distinguishing otherwise homophonous derivations from different words;

(21) It. *botta* 'stroke' → dim. *bott-ar-ella* vs. *bott-ic-ella* ← *botte* 'cask'; *cotta* 'infatuation' → *cott-ar-ella* vs. *cotta* 'surplice' → *cott-ic-ella*

In (22) the distinctive factor is the presence or absence of interfixes:

(22) *sole* 'sun' → *sol-ic-ino* vs. *solo* 'alone' → N *sol-ino* 'collar type'; *corto* 'short' → *cort-ina* vs. *corte* 'court' → *cort-ic-ina*; *riso* 'laugh' → *ris-ol-ino* vs. *riso* 'rice' → *ris-ino*

In (23) the presence or absence of interfixes distinguishes between diminutives and homophonous adjectival derivatives of the same bases:

(23) *carne* 'meat' → dim *carn-ina* vs. adj. *carn-ic-ino*
volpe 'fox' *volp-ic-ino* *volp-ino*

The only conceivable English examples seems to be *palace* → *palat-ial* vs. *palate* → *palat-al*.

3.9. Ninth criterion: Interfixes may fluctuate among each other between the same base and the same suffix:

(24) Sp. *café* → dim. *cafe-l-ito* = *cafe-t-ito*, R. *Glink-in/ov-skij*, Pol. *names* *X* → adj. *X-ow-ski* = *X-eń-ski*, It. *camp-(ic/er)-ello*, *grott-ic/er-ella*, cf. (2, 16).

No case of fluctuation found in English.

4. So, we may conclude that English has no antesuffixal interfixes, because the presumptive monophonemic or biphonemic elements lack sufficient free recurrence (criterion [2], including mutual interchangeability [9], cf. [6, 7, 8] and productivity ([3]), they are insensitive to brevity of the stem ([5], cf. Dressler 1984, 1986). Rather they are (nearly always predictable) parts of stem or suffix variants [2, 4].

5. Italian has interfixes similar to Spanish, Russian and Polish, even if fewer in quantity. And Italian interfixes may have some connotative meaning, e.g.

(25) *uomo* 'man' → dim. *om-ino* vs. *om-ar-ino*: pejorative
topo 'mouse' *top-ino* vs. *top-ol-ino*: more attractive
pianta 'plant' *piant-ina* vs. *piant-ic-ina*: more affective

Interfixed diminutives are different from double diminutives in that double diminutives may operate further diminution, whereas interfixes can only add connotation to the meaning of the simple diminutives:

(26) *casa* → *cas-etta* → *cas-ett-ina*, *libro* 'book' → *libr-ett-ino* 'very small book' # *libr-ic-ino*

Interfixation allows a connotative interpretation which, for example, can be anaphorically realized within a coreferential chain: In (27), *bach-er-ozzo* refers anaphorically to the preceding *baco* or dim. *bachino*, but gives a connotative reinterpretation of the same referent. The same occurs in (28). It would be very strange to use the reverse order:

(27) *Oh, c'è il baco/bach-ino anche in questa*! (after a pause) *Ah, il bach-er-ozzo si è servito a dovere*!
'Oh, there is a worm/small worm in this one! (while cutting an apple). Ah the little creature has helped itself!'

(28) A: *Vuoi un sors(in)o del mio caffè?* —
B: *Sìii!* — A: *Ehi, basta!*
'A: Want a (little) sip of my coffee? — B: Yes!!! A: *Hey, that's enough!*'
B: *Dài, te ne ho preso solo un sors-er-ello de niente!*
'Come on, I've only had a tiny sip'

Theoretical morphology

6. Now let's move from description to theory: Spanish, Russian, and Polish interfixes prove the existence of completely meaningless morphemes and thus disprove the close link between morphosemantics and morphotactics as postulated by structuralist morphology (cf. Aronoff 1976; Beard 1981). Italian interfixation has not arrived at this stage, since Italian interfixes may have connotative meaning.

As to the rule format of interfixation rules, Szymanek (1985a, b) is quite right that generative morphology — or any other process model — must allow for morphological epenthesis/insertion rules. This rule format is needed anyway for infixation rules. What is more interesting, is that interfixation rules must apply after the respective suffixation rules. E.g. if we anticipate certain facts from the last section on Italian interfixation (see **11**), then we observe 1) that the productivity of interfixes is much smaller than of the respective suffixes, 2) that the choice of an interfix depends on the choice of the suffix. We may add 3) that we have just seen that interfixed diminutive formations presuppose non-interfixed ones in coreferential chains. And 4) so far we have found that bases of interfixed and non-interfixed diminutives have the same distribution within word classes and semantic fields. All these facts are easily explainable if we assume that interfixation rules follow the respective suffixation rules. On the other hand, the Italian interfixes *-ar/er-* and -i[č]- are dissimilatorily blocked from applying after stem-final *-r-* and -[č] respectively (see **11.** (3)) i.e. the interfixation rule which applies after the suffixation rule must look back to the form of the base of the suffixation rule. Thus interfixation rules may apply under global conditions and may violate the Adjacency Condition (cf. Scalise 1984, 169—178).

7. But as adherents of Natural Morphology we are more interested in problems for our model. In universal markedness theory as the first subtheory of the model of Natural Morphology[7] interfixes are highly marked or

[7] Cf. Mayerthaler 1981, Wurzel 1984, Dressler 1982, 1985a, b, Dressler, Mayerthaler, Panagl and Wurzel (1987); and cf. the special volume on Natural Approaches to Morphology, in: Studia Grammatyczne 7 (1985).

very unnatural on two universal parameters: On the parameter of morphosemantic transparency they are totally opaque because they do not contribute anything to the meaning of the whole complex words that contain them. And on the parameter of morphotactic transparency or better, transparency of the signans, interfixes are very opaque because they are difficult to perceive because their typical shapes are a rime, a coda, or a rime of one syllable and an onset of the following syllable, whereas the most transparent shape of the morpheme consists of a syllable so that both the initial and the final morpheme boundary coincides with a syllable boundary.[8] Now if an interfix, such as all Italian interfixes, consists of rime and onset, no morpheme boundary coincides with a syllable boundary.

Thus interfixes are difficult to perceive. But is this a great disadvantage since they have no meaning anyway? Why bother to process them? In other words, we have a consequence of a third universal parameter, the parameter of diagrammaticity: Due to this parameter there is a tendency for morphosemantically transparent affixes to be morphotactically transparent as well. In this way we can explain why affixations with a compositional meaning represent usually productive affixation of a formally transparent nature.

And here morphosemantically opaque inerfixation tends to be morphotactically opaque as well. This represents a certain alleviation of the otherwise extremely marked character of interfixation.

Still interfixation is much more marked than suffixation or prefixation, but less marked than infixation. Therefore we can predict that cross-linguistically interfixation should be much rarer than suffixation or prefixation. And of course this prediction is empirically very well supported.

So far we have not found good examples of interfixes between prefixes and stems, first because the languages investigated are largely suffixing languages. However, due to psychological principles of processing (cf. Cutler et al. 1985) prefixation tends to be much more morphotactically transparent than suffixation, so that an avoidance of opaque interfixes would fit. Thus (both interradical and antesuffixal) interfixes seem (so far) to be a marginal sub-class of suffixes.

8. If we move to the typological subtheory[9] of Natural Morphology we can make e.g. the following prediction:

Premise I: If a universal naturalness parameter is higher valued in language type X than in language type Y, then morphological phenomena which are very unnatural on this parameter should occur much less in type X than in type Y.

[8] On morphemes and syllables as processing units see now Norris and Cutler 1985.
[9] cf. Dressler 1985a, b, c; Skalička 1979.

Premise II (as identified above): Interfixation is very unnatural on the parameters of morphosemantic/tactic transparency.

Premise III (constitutive of the typological subtheory): The parameters of morphosemantic/tactic transparency are much higher valued in the agglutinating type than in the inflecting type.

Conclusion and Prediction: The more agglutinating a language or a subpart of morphology is the less probable it is that (interradical/antesuffixal) interfixation occurs.

This prediction is empirically well supported by the distribution of interradical and antesuffixal interfixation combined: Agglutinating languages such as Turkish and Hungarian do not have interfixes, highly inflecting languages such as Russian, Polish, Sanskrit, Ancient Greek have, also moderately inflecting languages such as German, Dutch, the North Germanic, the Romance languages, or Estonian which has changed from an agglutinating to an inflecting language. English represents the inflecting type only in its Latinate word formation; and it's precisely there that English has interradical interfixation.

One intervening variable is 'morphological richness' which favours the occurrence of morphological phenomena: Germanic languages are relatively richer in composition: they have interfixation just there. Slavic and Romance languages are richer in derivational word formation: they have both interradical and antesuffixal interfixation (with the exception of French which has only interradical interfixation). However neither the universal nor the typological subtheory of Natural Morphology can explain so far why Spanish has more antesuffixal interfixation than Italian. These are language specific problems with which we deal elsewhere. Also the question why introflecting languages have no interfixation has been dealt with elsewhere (Dressler 1984). This may show the explanatory range of Natural Morphology in the area of interfixation.

9. As to diachronic explanation, Malkiel (1958 etc.), Dougherty (1984), Dressler (1986) deal with the development of antesuffixal interfixes.[10] In regard to decay and loss of interfixes we may mention French which earlier must have had interfixes (cf. the comparative-diachronic equation Fr. *fort-er-esse* = Sp. *fort-al-eza* = It. *fort-ezza* without interfix), but lost them as the Romance language where the inflecting type (which admits interfixes, see above) is most weakly represented.

[10] For Italian see our forthcoming paper at the Bologna meeting of the Società di Linguistica Italiana.

Towards Morphopragmatics

10. Authors dealing with pragmatic aspects of word formation rules usually think of questions of denotative reference in lexical usage (norms), e.g. as often mentioned by E. Coseriu why a German *Eis#verkäufer* sells ice (object), but a *Straß+en#verkäufer* does not sell streets, but rather on the streets/roads, cf. for other models of dealing with similar questions: Bauer (1979), Anwar (1984), Beard (1978), Romaine (1983). Such reflections can tell us something about the existence of actual words with actual lexical meanings and thus about the frequency of derivations of a given Word formation rule. But this does not help very much for the question of potential/possible words, the proper domain of productivity studies in word formation (cf. Aronoff 1976, 1980, 1983), e.g. an *Eis#verkäufer* might be someone who sells something to ice-skaters while skating "on ice", and a *Straß+en#verkäufer* may be someone who sells roads (road building) to Saudi Arabia (as has been the case e.g. with Austrian road-builders). Therefore all productivity tests which do not start with nonce-words as bases of word formation rules are marred by the problem that subjects are influenced by cognitive presuppositions about possible cognitive worlds in which objects may exist which could be labelled by non-existing derivations of a word formation rule.

On the other hand we know that working with nonce-forms as bases of word formation rules introduces an artefact of a laboratory experiment which increases the difficulty of this metalinguistic task for the subjects.

Studying the productivity of meaningless interfixes offers the advantage of largely avoiding both problems mentioned so far: 1) we can use existing words as bases of the word formation rules to be studied, 2) we can study only the question whether an interfix may be inserted into a complex word which has no interfix. This interfixless word may be either an actual word or only a potential word. In any case the cognitive referential presuppositions regard the suffixation rule, but not the interfixation rule, or it regards in Robert Beard's model, the morphosemantic derivation rule, but not the morphotactic interfixation rule which presumably has no morphosemantic correspondent of its own. Thus, in studying the productivity of interfixes, we can largely avoid the problems of reference in productivity tests.

11. The elements of a probability model for Italian interfixes that precede diminutive suffixes are tentatively the following:

Among the structural factors there is

1) just one absolute constraint: no interfix may be inserted before the diminutive suffix *-etto*.

2) A default factor is: no base should have 3 or more syllables, see **3.5.**

3) There is a dissimilatory blocking (cf. Hjordt-Vetlesen 1981 for Rumanian) of interfixes by the stem-final consonant (not of the pre-final syllable), as table (29) shows:

(29) dissimilatory stem-final C	blocking: interfix	instances without interfix interf. impossible
-č-	*-ič-*	6
-r	*-ar/er*	16
(pre-final-syll. *-r-*)	*-ar/er*	(4)
-l-	*-ol-*	5

	instances with interfix	instances where an interfix would be possible, but is not attested
-č-	0	2
-r-	0	2: *scon^{s}tre+e^{s}r+e^{s}llo*
(-r-)	(1): *cor^{s} s+e^{s}r+e^{s}lla*	(21): *tor^{s}t+e^{s}r+e^{s}lla*
-l-	4: *fil+ol+ino, pel/grill giall+ol+ino*	

4) The frequency (rank order) of attested interfix-suffix-combinations is:

(30) interfix-suffix-combinations: *-a/er+ello* 199 (33%) > *-ol+ino* 137 (24%) > *-ic+ino* 134 (23%) > *-ic+ello* 92 (16%) > *ici+a/ott+olo* 6 > *-er+ino* 4

5) This can be transformed into the following if-then statements.

(31) a) if *-ello*, then 68% *-a/er-* > 32% *-ič-*
b) if *-ino*, then 50% *-ol-* > 49% *-ič-* > 1% *-er-*

These frequency counts are based on all attested interfixed diminutive derivations including all lexicalizations. A series of tests with informants where they either have to produce their own nonce forms or to evaluate nonce forms presented to them will help to induce probability factors.

6) Note that only 7% of all *-ino* diminutives are interfixed forms in *-ic-ino*, and only 1,6% of all *-ello* diminutives are forms in *-ic-ello*.

7) For the position of an interfixed diminutive formation in an anaphoric coreference chain, see above §5.

8) The tropic use of interfixed diminutive formation consists in dedramatizing facts, e.g.

(32) *il nostro scontr+ic+ino quotidiano*
'our daily nice little car-crash'

(33) *È veramente un pass+er+ello avanti!*
'(sc. Your theory) is really a nice little step forwards!'

9) Affective connotation is most clearly visible in the so-called "diminutiva infantilia" (Staverman 1953). They are most often used in baby talk, less often in talk to children in general, least in puristic talk, because purist norms proscribe the use of such diminutives, especially of double and interfixed diminutives, as a sign of affection. Factors involved in the use of interfixed diminutives with other connotative meanings are more difficult to assess; even the distinction and classification of different connotative meanings is not obvious at all: e.g. Staverman (1953) differentiates diminutiva socialia sive familiaria (even: culinaria), modesta (see above 8)), puerilia, ludicra, and infantilia sive inepta.

10) Sociolinguistic factors are interdependent variables: e.g. sex (women seem to use more interfixed diminutives than men) interacts with factor 9), with age (middle-aged persond seem most inclined to use diminutives), and with level of education (which affects influence of puristic prescription and proscriptions, see factor 9)).

11) The most elusive factor is the linguistic and metalinguistic capacity, as shown in production and evaluation tests. Here differences of pragmatic imagination are included, which hopefully refer rather to suffixation rules than interfixation rules.

Conclusion

12. In the descriptive part we have presented criteria which allow to decide whether a language has prototypical antesuffixal interfixes (Spanish, Russian, Polish — but also other Slavic and Romance languages), less typical ones (Italian) or only stem- and suffix variants (English). In the theoretical part we have discussed problems that interfixes pose to Generative Morphology and shown the explanatory range of Natural Morphology in regard to interfixes at the present stage of its development.

Finally, in a more programmatic part we have enumerated factors determining the application of Italian interfixation rules in diminutive formation. These factors are easy to operationalize and thus allow morphopragmatics to become a testable, empirical part of word formation studies. In this way a scientific "linguistique de la parole" becomes possible within the linguistic (not only psycholinguistic) study of word formation.

References

Anwar, M. S.: Lexical innovation as a pragmatic analysis of symbols. In: Martin, P. (ed.): Actes du 10 Colloque, Societé Internationale de Linguistique Fonctionelle. Université Laval 1984, 81—84.
Aronoff, M.: Word Formation in Generative Grammar. MIT-Press, Cambridge 1976.
Aronoff, M.: Potential words, actual words, productivity and frequency. In: PICL 13 (1983), 163—171.
Bauer, L.: On the need for pragmatics in the study of nominal compounding. In: Journal of Pragmatics 3 (1979), 45—50.
Beard, R.: Derivations within derivations. In: Birnbaum, H. (ed.): American Contributions to the 8th International Congress of Slavicists I. Columbus, Slavia (1978), 126—154.
Beard, R.: The Indo-European Lexicon. Amsterdam, North Holland 1981.
Cutler, A. et al.: The suffixing preference: a processing explanation. In: Linguistics 23 (1985), 723—758.
Dougherty, P.: Wortbildung mit dem Segment *-ov-* in der modernen russischen Schriftsprache. Ph. D. Thesis, University of Vienna.
Dressler, W. U.: On word formation in Natural Morphology. In: Wiener linguistische Gazette 26 (1982) 3—14 (= PICL 13, 172—182).
Dressler, W. U.: Zur Wertung der Interfixe in einer semiotischen Theorie der Natürlichen Morphologie. In: Wiener Slawistischer Almanach 13 (1984), 33—45.
Dressler, W. U.: Morphology. Ann Arbor, Karoma Press 1985a.
Dressler, W. U.: On the predictiveness of Natural Morphology. In: Journal of Linguistics 21 (1985b), 321—337.
Dressler, W. U.: Typological Aspects of Natural Morphology. In: Wiener linguistische Gazetta 35—36 (1985c), 3—26.
Dressler, W. U.: Forma y función de los interfijos. In: Revista Española de Lingüística 16 (1986), 381—395.
Dressler, W. U., Mayerthaler, W., Panagl, O. and Wurzel, W. U.: Leitmotifs in Natural Morphology. Amsterdam, Benjamins (1987).
Jaeggli, O.: Spanish Diminutives. In: Niessel, F. H.: Contemporary Studies in Romance Languages. Indiana University Linguistics Club, Bloomington 1980, 142—158.
Lehnert, M.: Reverse Dictionary of Present-Day-English. VEB Enzyklopädie, Leipzig 1971.
Lopatin, V.: Tak nazyvaemaja interfiksacija i problemy struktury slova v russkom jazyke. In: VJa 4 (1975), 24—37.
Malkiel, Y.: Los interfijos hispánicos. In: Miscelánea homenaje A. Martinet II. Gredos, Madrid 1966, 305—364.
Norris, D. and Cutler, A.: Juncture detection. In: Linguistics 23 (1985) 689—705.
Prati, A.: Antesuffissi. In: Italia Dialettale 18 (1942), 75—166.
Romaine, S.: On the productivity of word formation rules and limits of variability in the lexicon. In: Australian Journal of Linguistics 3 (1983), 177—200.
Scalise, S.: Generative Morphology. Foris, Dordrecht 1984.
Skalička, V.: Typologische Studien. Vieweg, Braunschweig 1979.
Staverman, W. H.: Diminutivis Neerlandica. In: De Gids 116 (1953), 407—419 (here cited after: van Zonneveld, R.: Affix-Grammatika. Groningen 1983, 161—162).
Szymanek, B.: English and Polish Adjectives. Katolicki Universytet Lubelski, Lublin 1985a.
Szymanek, B.: On intermorphic extensions in English and Polish. In: Gussman, E. (ed.): Phono-Morphology. 1985b, 177—191.
Tekavčić, P.: Sull' analisi morfematica di un tipo di derivati italiani. Un problema dei cosiddetti interfissi. In: Studia romanica et anglica Zagrebiensia 25/26 (1968), 69—85.
Wurzel, W. U.: Flexionsmorphologie und Natürlichkeit. In: Studia Grammatica 21 (1984).
Zemskaja, E. A.: K probleme množestvennosti morfologičeskich interpretacij. In: Zemskaja, E. A. (ed.): Razvitie sovremennogo russkogo jazyka. Nauka, Moscow 1972, 69—88.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 69–91 (1986)

PREDICATE FORMATION IN FUNCTIONAL GRAMMAR

C. DE GROOT
(University of Amsterdam, The Netherlands)

Introduction

In natural languages we often find systematic relations between several classes of predicates (verbs, adjectives and nouns) such as intransitives, transitives, causatives, inchoatives, reflexives etc. This is for instance the case in Hungarian, consider:

(1) (a) *szép* 'pretty'
(b) *szépül* 'get prettier'
(c) *szépít* 'make pretty'
(2) (a) *mos* 'wash'
(b) *mosakodik* 'wash oneself'
(c) *mosat* 'have wash'

An account of the relation between these types of predicates merely in terms of "stem/root + affix" is, of course, not sufficient, because it does not do justice to relevant properties such as valency as well as to the relation between the predicates and the states of affairs which they can designate. The first aim of this paper is to demonstrate the relevance of valency and states of affairs in derivational processes. The second aim is to demonstrate that the theory of Functional Grammar (FG) as developed by Dik (1978)[1] offers a formalism which can account for all these properties.

In this paper I shall make use of material and views presented in earlier writing on predicate formation in FG. Although no efforts have been made to give a complete survey of types of predicate formation rules and the effects they can have on their input, the paper can be seen as offering a "state of the art" description of predicate formation.

The first section presents an outline of the theory of Functional Grammar, followed by some remarks on productivity in section **2**, and the place of morphology in FG in section **3**. States of affairs, properties of predicates

[1] And further developed in Dik (1980), Dik ed. (1983, 1985), Hoekstra et al. eds. (1981), Bolkestein et al. (1981) and Bolkestein et al. eds.

and the semantic functions of their arguments are dealt with in section 4. The last section discusses a number of predicate formation rules relevant to Hungarian.

1. Functional Grammar

FG aims at a maximum of practical applicability in the analysis of diverse aspects of language and language use. An attempt is made to reach this goal by (i) maximizing the degree of typological adquacy, while (i) minimizing the degree of abstractness of linguistic analysis. By degree of abstractness is meant the distance (as measured in terms of rules, operations, or procedures) between the structures postulated for a given language on the basis of the theory, and the actual linguistic expressions of that language which are reconstructed in terms of these structures. Constraints restricting the degree of abstractness are:

(i) transformations in the sense of structure-changing operations are avoided;
(ii) empty elements in underlying structure which do not receive expression are avoided;
(iii) filter devices are disallowed;
(iv) abstract lexical decomposition is not applied (instead the semantic relations between words are accounted for through meaning definitions).

The overall layout of FG can be indicated globally as follows:

(i) the f u n d, which consists of a set of predicates and a set of t e r m s (including those predicates and terms that are derived by formation rules);
(ii) the p r e d i c a t i o n s, which are structures created by combining predicates and terms;
(iii) e x p r e s s i o n r u l e s, which map predications onto linguistic expressions.

(i) The fund contains a l e x i c o n, i.e. a list of all predicates, or contentives, of a language. They are called b a s i c p r e d i c a t e s. The set of basic predicates can be extended with a set of d e r i v e d p r e d i c a t e s by means of a system of productive p r e d i c a t e f o r m a t i o n r u l e s, such as rules of derivation and composition. All other formations of a basic predicate which cannot be considered the result of some productive rule (this also includes non-productive aspects of inflection) are given in the lexicon, too (see section 2 for further discussion).

Predicates are expressions designating properties or relations. They are contained in p r e d i c a t e - f r a m e s, structures which specify their fundamental semantic and syntactic properties, such as (i) the syntactic category of the predicate (Verbal, Nominal, Adjectival), (ii) the number of arguments, (iii) the semantic functions of the arguments (Agent, Goal, Recipient etc.). Consider the following example:

(3) $\text{give}_{V}\ (x_1)_{Ag}\ (x_2)_{Go}\ (x_3)_{Rec}$

The order in which the predicate and the arguments are given has no direct or necessary relation to the linear order in which these constituents will finally be realised. Predicate-frame (3) could just as well be given in another linear form or in a two or threedimensional form. The representation of predicate-frame (3) is purely a matter of convention.

Basic and derived predicate-frames are together referred to as n u c l e a r p r e d i c a t e - f r a m e s. All predicate-frames have a m e a n i n g d e f i n i t i o n, for instance:

(4) $\underset{df}{\begin{matrix}\text{boy}_{N}\ (x_1)_{\emptyset} \\ \text{child}_{N}\ (x_1)_{\emptyset}: \text{male}_{A}\ (x_1)_{\emptyset}\end{matrix}}$

Nuclear predicate-frames can be extended by s a t e l l i t e s (non-arguments). The semantic functions of arguments express the relations between the predicate and the arguments; the semantic functions of satellites express the relation between the s t a t e o f a f f a i r s (designated by the predicate-frame) and the stellites. Consider:

(5) $[\text{buy}_{V}\ (x_1)_{Ag}\ (x_2)_{Go}\ ]_{ACTION}\ (y_1)_{Loc}$

The variables indicating the arguments in predicate-frames and satellites can be replaced by inserting t e r m s, i.e. the forms underlying referring expressions.[2] If such insertion is applied to all open slots of a given predicate-frame, the result is a (closed) p r e d i c a t i o n. Consider:

[2] Two types of terms are distinguished: (i) basic terms, expressions which can only function as terms and are given as such in the lexicon (e.g. personal pronouns, proper nouns, question words) and (ii) derived terms, which can be formed by the following general schema:

(i) $(\Omega x_i: \Phi_1(x_i): \Phi_2(x_i): \ldots : \Phi_n(x_i))$

Here x_i is the term variable symbolizing the intended referent of the term; the symbol Ω indicates one or more term operators (operators for definiteness, number etc.); each $\Phi(x_i)$ indicates some 'open predication in x_i', that is, a predicate-frame all of whose argument positions have been bound except for x_i. Each open predication in x_i can be regarded as a restrictor specifying some property which x*i* must have in order to qualify as a potential referent of the term. Restrictors are stacked onto each other through the relation indicated by ':' ('such that').

(6) [buy_V (Peter)$_{Ag}$ (a new coat)$_{Go}$] (the market)$_{Loc}$

Many grammatical elements, such as those expressing Tense and Aspect distinctions, are introduced by means of operators:[3]

(7) Past buy_V (Peter)$_{Ag}$ (a new coat)$_{Go}$

(ii) Alongside the semantic functions given in the predications themselves, there are also two other types of functions, syntactic and pragmatic. Funtional Grammar recognizes only two s y n t a c t i c f u n c t i o n s: Subject and Object. Syntactic functions express the perspective from which a certain state of affairs is presented. Different syntactic function assignment accounts for the difference in the expression of the same state of affairs in (8a—b).

(8) (a) Past buy_V (Peter)$_{AgSubj}$ (a new coat)$_{GoObj}$
'Peter bought a new coat'
(b) Past buy_V (Peter)$_{Ag}$ (a new coat)$_{GoSubj}$
'A new coat was bought by Peter'

In (8a), the state of affairs is presented from the point of view of the Agent; in (8b) it is presented from the point of view of the Goal. These differences are formally captured by assigning the Subject function to the Agent or the Goal of the underlying predication.

In FG, four P r a g m a t i c f u n c t i o n s are distinguished. There are two pragmatic functions external to the predication proper, Theme and Tail, and two pragmatic functions internal to it, Topic and Focus, Consider:[4]

(9) That new coat, he bought it one the market, Peter
Theme Topic Topic Focus Tail
[————————Predication————————]

[3] Operators are applied at several levels of the predications. A distinction is made between predication operators (e.g. illocution), predicate operators (e.g. tense, aspect, mood) and term operators (e.g. definiteness and number). Operators typically belong to subsystems with a limited number of 'values' from which a choice can be made. They reflect what is more traditionally known as the grammatical, morpho-syntactic, or morphosemantic 'categories' relevant to a language.

[4] The Theme specifies the universe of discourse with respect to which the subsequent predication is presented as relevant; the Tail presents, as an 'afterthought' to the predication, information meant to clarify or modify it. The Topic one of the two pragmatic functions internal to the predication proper, present the entity/entities 'about' which the predication predicates something in a given setting ('he' and 'it' in (9)). The other function, Focus, presents what is relatively the most important or salient information in a given setting ('on the market' in (9)).

(iii) The expression rules form the last component in the model. The expression rules determine the way in which functional structures are mapped onto the syntactic structures of linguistic expressions. This component takes care of constituent ordering, case marking, voice, copula support, auxiliary elements, agreement etc.

The organization of a Functional Grammar is given in figure 1 below.

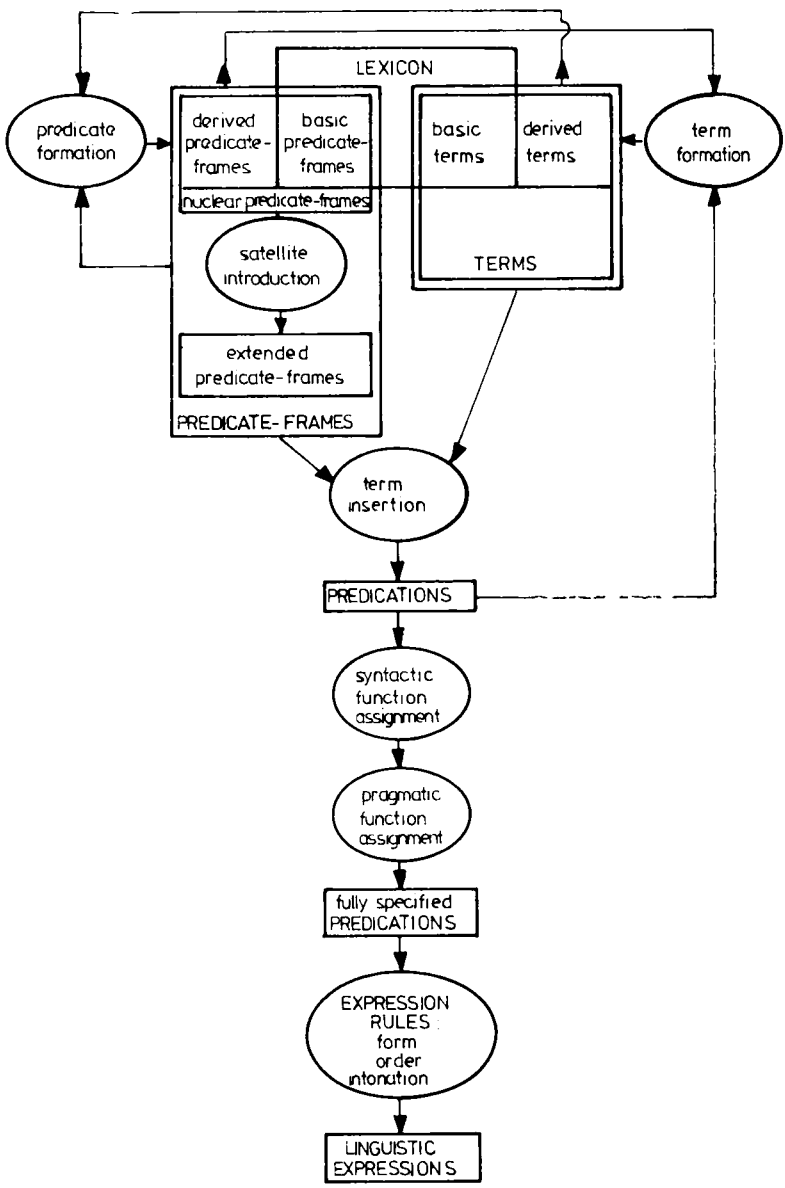


Fig. 1. The organization of a Functional Grammar

2. Productivity, rules and regularities

All rules, such as predicate formation rules and expression rules, are considered to be completely productive, where productivity is defined in terms of the ability of a competent speaker to apply the process in question in correctly deriving output expressions which he may never have heard before. Examples of rules are for instance the application of the causative formative suffix *-tat/-tet* in Hungarian, or the past tense suffix *-ed* in English. In a lexicon of Hungarian we will find entries such as *sétál* 'walk' and *köhög* 'cough'. In case of causativization a predicate formation rule will introduce the formative suffix *-tat/-tet*: *sétáltat* 'make walk', *köhögtet* 'make cough'. In a lexicon of English we will find entries such as *walk* and *cough*. When the past tense has been applied an expression rule will take care of the application of the past tense suffix *-ed* to the verbal stem: *walked* and *coughed*.

All formations of a basic predicate which cannot be considered the result of some productive rule are given in the lexicon. Consider *buy* opposed to *walk* (Dik 1979):

(10) (a) $\{\text{walk}_V\}$ $(x_1)_{Ag}$
(b) $\{\text{buy}_V, \text{Past bought}_V, \text{Past Part bought}_V\}$ $(x_1)_{Ag}$ $(x_2)_{Go}$

Patterns such as *sing, sang, sung* and *ring, rang, rung* are stored in the lexicon as such (cf. (10)b). Their common patterning constitutes a regularity in the lexicon.

3. Morphology

In the overall structure of FG, as given in figure 1, there is no mention of a morphological component as such. However, the location of morphology in the model is well defined.

Morphemes can be divided into two categories. The first category consists of an open class of elements, i.e. there is no limit to the number of elements. The second category consists of a closed class of elements, i.e. there is a limit to the number of elements. Examples of elements which belong to the open category are *work, walk, train, table, nice,* and *long*. Elements which belong to the closed category are for instance *the, of, re-, -ed,* and *-s*. Elements of the open category can be characterized as l e x i c a l e l e m e n t s, and elements of the closed category as g r a m m a t i c a l e l e m e n t s. FG reflects this distinction in the following way: lexical elements are contained in the lexicon, grammatical elements (which include derivational and inflectional affixes) in

other components. We can now say that the domain in which morphology is located can be anywhere but the lexicon.[5]

The basic distinction between derivational and inflectional morphology can also be considered to be a theory-internal one, as has been pointed out by Watters (1985). Given the provision that the lexicon does not belong to the domain in which morphological processes apply, we may characterize morphology in FG with a slight modification after Watters (1985) in the following fashion:

(11) (i) Derivational morphology is what is relevant to the predicate formation rules
(ii) Inflectional morphology is what is relevant to the expression rules

FG provides a natural, theory-internal distinction between derivational and inflectional morphology. Derivational morphology is that morphology found in the predicate formation component and inflectional morphology is that morphology found in the expression rules. Note that predicate formation rules and expression rules do not necessarily involve any morphology. This is for instance the case in the formation of certain derived intransitive verbs in English (e.g. intransitive *write* 'this pen writes well' from transitive *write*), or in the expression of the order of constituents. Thus, it is incorrect to say that predicate formation and expression rules are synonyms of respectively derivational and inflectional morphology.

The following figure indicates the location of morphology in FG:

Morphology	open class of elements	closed class of elemets	
		derivational elements	inflectional elements
Functional Grammar	lexicon	predicate formation rules	expression rules

Fig. 2. FG and morphology

As such, morphology within FG is diffuse, not being localized in any one part of the theory. It should be noted that the two types of morphology are distinct in terms of their ordering within complex words. This ordering can be specified by the following schema, which by and large holds across languages (Watters 1985, Bybee 1985):

(12) [inflection [derivation [stem/root] derivation] inflection]

[5] One example of morphology in the lexicon may be the following. Regularities in the lexicon will usually go unnoticed, but may occasionally be 'abstracted' from the relevant forms by the speaker, and then lead to incidental innovations of the *bring — brang — brung* type (cf. Dik in prep. ch. 10).

4. Predicates

4.1. States of affairs

Nuclear predications consist of predicates and terms. Terms refer to entities in some world, and predication designate properties of, or relations between such entities. A nuclear predication as a whole designates a set of states of affairs. The term state of affairs is used in the broad sense of "conception of something which can be the case in some world".

States of affairs can be divided into different types, according to the values which they can have for a number of distinguishing parameters. These parameters and their different values together define a semantic cross-classification of states of affairs.[6]

The most important semantic parameters defining the typology of states of affairs are given in (13); one test-frame is given for each of the parameters in (14) through (17):

(13) +/— Dynamic [dyn]
+/— Momentaneous [mom]
+/— Control [con]
+/— Telic [tel]

(14) Dynamic

A [—dyn] state of affairs is a state of affairs which does not involve any change. One criterion for distinguishing between [+dyn] and [—dyn] states of affairs is that the latter do, but the former do not combine with stallites of Speed:

(a) *walk slowly* [+dyn]
(b) **stand slowly* [—dyn]

(15) Momentaneous

A test which distinguishes between [+mom] and [—mom] states of affairs is the so-called "almost-test". Consider the following two sentences:

(a) *John almost reached the summit.* [+mom]
(b) *John almost read a book.* [—mom]

The first sentence tells us that John did not reach the summit. The second example is ambiguous in the following fashion:

'John intended to read a book but changed his mind and did nothing at all'

'John began to read a book and he almost but not quite finished reading it'

[6] For discussion and other references see Bridgen (1984), Dik (1978), De Groot (1983, 1985), Vester (1983), and Vet (1980).

(16) Control

A test which distinguishes between [+con] and [−con] states of affairs is based upon the consideration that [+con] states of affairs can occur in the true imperative, whereas the other type cannot. For instance:

(a) *go*! [+con]
(b) **know*! [−con]

(17) Telic

Telic states of affairs can be extended with the 'within an hour' phrase, atelic states of affairs cannot.[7] Consider:

(a) *reach x within an hour* [+tel]
(b) **work within an hour* [−tel]

It seems, however, that not all combinations of parameters within one state of affairs are possible. SoAs cannot be both non-dynamic and telic, or both momentary and atelic. It seems that some parameters entail others (cf. Vester 1983, De Groot 1983). Consider:

(18) (a) [−dyn] > [−tel]
(b) [+mom] > [+tel]

4.2. Nuclear semantic functions

4.2.1. First argument

The structure of predicate-frames has no direct relevance for the final positions the constituents take in the syntactic structures of actual linguistic expressions. There is, however, another sort of ordering, which is hierarchical rather than linear. It concerns the intrinsic relationship among the semantic functions, such that some semantic functions are more 'central' to the predicate than others. If only one argument is involved in an Action state of affairs ([+con], [+dyn]), that argument will necessarily designate the entity controlling the Action. The arguments having this property are assigned the semantic function of A g e n t. If only one argument is involved in a dynamic non-controlled state of affairs, a Process, that argument will designate either the entity that is primarily involved in a Process, or the non-controlling entity instigating a Process. The arguments having these properties are assigned the semantic functions of P r o c e s s e d and F o r c e respective. As for

[7] I interpret the notion Telicity in the sense of Comrie (1976) in the following way: if a state of affairs has built into it a terminal point, the state of affairs has the feature Telic.

non-dynamic states of affairs, the following two semantic functions are recognized: Positioner ([+con], [−dyn]) and Zero ([−con], [−dyn]).

Thus any one-place predicate has an argument with either Agent, Processed, Force, Positioner, or Zero function. Because we often want to be able to say that a certain grammatical rule applies to any first argument, no matter what its semantic function is, the notion first argument has been introduced (De Groot 1981).[8] The first argument of a predication is defined as the most central (possibly the only) argument of the predication.

4.2.2. Non-first arguments

Apart from the semantic functions of first arguments, at least the following further semantic functions must be distinguished as potentially relevant for nuclear predicate-frames:

(19) Goal (Patient): the entity affected or effected by the operation of some Controller (Agent or Positioner) or some Force.
Recipient: the entity to which something is transferred.
Direction: the entity towards which something moves/is moved.
Source: the entity from which something moves/is moved.
Location: the place where some entity is located.

The following constructions exemplify these functions, and some of their possible combinations:

(20) Peter_{Ag} gave the book_{Go} to Mary_{Rec}
(21) Mary_{Ag} sent the boy_{Go} to the $\text{office}_{\text{Dir}}$
(22) $\text{Thomas}_{\text{Ag}}$ read the $\text{newspaper}_{\text{Go}}$
(23) Ann_{Ag} returned from $\text{England}_{\text{So}}$
(24) John_{Pos} lives in $\text{London}_{\text{Loc}}$
(25) $\text{Vera}_{\text{Proc}}$ received a $\text{letter}_{\text{Go}}$ from her aunt_{So}

In contrast to what is possible with respect to first arguments, it is not possible to generalize over a group of semantic functions as "second argument" or as "third argument". This is because, apart from the functions of Goal and Recipient, all functions given under (19) also occur as semantic functions of satellites. For instance:

[8] The notion of First argument is not the same as the notion of External argument (Williams 1981) since FG does not recognize 'VP'.

(26) Thomas read the newspaper (in the garden$_{Loc}$)
(27) Jane prepared a supper (from left-overs$_{So}$)

In other words, there is no defined group of semantic functions which exclusively relates to second arguments or third arguments.

Goal and Recipient function apply only to arguments and not to satellites. Prototypically, Goal relates to the second argument and Recipient to the third argument of a predicate. It seems that the second argument in three-place predicates always has the function of Goal. Consider the following constructions:

(28) (a) *John spread the butter on the bread.*
(b) *John spread the bread with butter.*
(29) (a) *Mary taught the children to sing.*
(b) *Mary taught geography to the children.*

It has been argued that verbs such as *spread* (Dik 1980) and *teach* (Workgroup 1981) must be assumed to have two different predicate-frames according as they occur in constructions corresponding to (28a) and (28b), or (29a) and (29b). Consider:

(30) (a) spread$_v$ (John)$_{Ag}$ (the butter)$_{Go}$ (the bread)
(b) spread$_v$ (John)$_{Ag}$ (the bread)$_{Go}$ (the butter)
(31) (a) teach$_v$ (Mary)$_{Ag}$ (the children)$_{Go}$ (to sing)
(b) teach$_v$ (Mary)$_{Ag}$ (geography)$_{Go}$ (the children)

We will return to these patterns in section **5.4.**

4.3. The relation between parameters and predicates

Predicate-frames designate sets of states of affairs. Which sets are designated is, however, partly determined by the predicates themselves. For instance, a predicate such as *walk* will only occur in dynamic states of affairs and not in non-dynamic states of affairs. In this light De Groot (1985) has argued that predicates can be characterized in terms of the parameters determining the typology of states of affairs. I shall refer to the specifications predicates have for the states of affairs in which they can occur as "features of predicates". Thus, we can say that the predicate *walk* has a feature [+dyn]. A distinction can be made between two types of features: (i) inherent features, [dyn] and [mom], and (ii) contingent features, [con] and [tel].

Predicates only occur either in dynamic states of affairs or in non-dynamic states of affairs, or either in momentary states of affairs or in non-

momentary states of affairs. We will therefore say that [dyn] and [mom] are inherent features of predicates. Consider:

(32) (a) *Peter is running.* [+dyn]
(b) *Mary is standing* [−dyn]
(33) (a) *Louis has hit Charles.* [+mom]
(b) *Joan is reading a book.* [−mom]

Contrary to what is said about predicates and the features [dyn] and [mom], there are predicates which can occur in both controlled and non-controlled states of affairs, or in both telic and atelic states of affairs. For instance:

(34) (a) *Mary stands in the corner.* [+con]
(b) *the cupboard stands in the corner.* [−con]
(35) (a) *John read the book.* [+tel]
(b) *John is reading.* [−tel]

The contingent features [con] and [tel] cannot be associated with lexical properties of the predicates (whereas inherent features can), but they can be associated with arguments or satellites of predicates. For instance, the feature [con] always affects the first argument of a predicate and not any other. We can say that [con] binds the first argument. Consider (36a), where John is the controller of an Action, and (36b) where Mary is a "non-controller" in a Process:

(36) (a) give_{v} $(\text{John})_{\text{Ag}}$ $(\text{the book})_{\text{Go}}$ $(\text{Mary})_{\text{Rec}}$ [+con]
(b) $\text{receive}_{\text{v}}$ $(\text{Mary})_{\text{Proc}}$ $(\text{the book})_{\text{Go}}$ $(\text{John})_{\text{So}}$ [−con]

Telicity, which is sometimes also described as "goal-orientedness", can be associated with those arguments or satellites of predicates which set the terminal point in the state of affairs. In general, telicity binds the affected argument of a predicate or the Directional argument/satellite in a predication. Consider the following examples, where the phrases in italics define the terminal point:

(37) (a) John read *the newspaper*$_{\text{Go}}$
(b) Mary walks *to the station*$_{\text{Dir}}$

Note, however, that not all predicates are compatible with both [+con] and [−con] or [+tel] and [−tel] states of affairs. Consider:

(38) (a) *exist* [−con]/*[+con]
(b) *reach* [+tel]/*[−tel]

In these cases we may also say that [con] and [tel] are inherent features. There remains the difference concerning the relation between the features and the predicate-frames: [dyn] and [mom] relate to lexical properties of predicates, [con] and [tel] relate to participants or entities involved in the state of affairs.

This leads to two constraints on the relation between the parameters defining the typology of states of affairs and certain properties of predicates:

(39) Predicate-frames cannot be specified for both (opposite) values of inherent features.

(40) Predicate-frames cannot be specified for features which bind different entities.

In relevant cases, I shall use the following notational convention for indicating the relation between features and predicate-frames:

(41) $[+\text{dyn}]$, $[-\text{mom}]$ read_{V} $([+\text{con}]\ x_1)_{\text{Ag}}$ $([\pm\text{tel}]\ x_2)_{\text{Go}}$

5. Predicate formation

5.1. Introduction

In section 1 a distinction was made between basic and derived predicates. The lexicon represents the stock of basic predicates which language users must know in order to be able to use them, while the predicate formation component reflects what they may form by themselves. Thus, derived predicates are those predicates which can be formed by means of some synchronically productive rule. All predicates, whether basic or derived, are contained in predicate-frames which specify their semantic properties. The input of a predicate formation rule can consist of basic and derived predicate-frames. The output predicate-frames of a predicate formation rule are necessarily derived. Predicate formation can schematically be represented as follows:

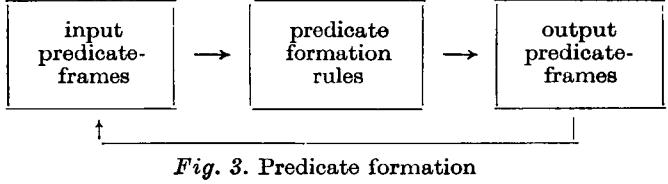


Fig. 3. Predicate formation

Differences between the input and the output of predicate formation rules can be given in two different ways. This can be done by describing the differences between the input and output predicate-frames in terms of

(i) valency, semantic functions and categories, and
(ii) the features [dyn], [mom], [con], and [tel].

Dik (1980) has argued that predicate formation rules may have different sorts of effects on the input predicate-frame. The most important effects are given in (42):

(42) (i) Effects on the valency of the predicate
(a) valency extension
(b) valency reduction
(ii) Other effects on the input predicate-frame
(c) semantic function shift of the arguments of the predicate
(d) semantic modification of predicate
(e) change in the syntactic category of the predicate

Given the theory concerning the relation between predicates and features of SoAs (see section **4.3.** above) we can postulate the following types of differences between input and output predicates:

(43) (a) opposite values of inherent features
(b) contingent features bind different entities

In the following sections I shall give examples of several types of predicate formation rules relevant to Hungarian. Since it is the function of these examples to illustrate some aspects of predicate formation, I shall not discuss these predicate formation rules with respect to productivity and special properties of the input and output predicate-frames. I shall confine myself to showing that some fundamental properties of predicate formation can easily be accounted for within the formalism offered in FG.

The predicate-frames in the examples will not be fully specified for semantic, syntactic and pragmatic distinctions. The Hungarian examples are meant merely to illustrate the valency of the predicates, the derivational markers, and the expression of the semantic functions.

5.2. Valency extension

A classic example of a derived construction involving valency extension is causative predicate formation. Many languages have a predicate formation rule which derives a causative predicate from a non-causative predicate.[9] In general such a rule can be formulated in the following way:

[9] See for instance Dik (1980, Ch. 3), Junger (1985 a), Schaaik (1985), and Vet (1985).

(44) input: pred_V (x_1) (x_2) . . . (x_n)
output: pred_V-E $(x_0)_{\text{Causer}}$ $(x_1)_{\text{Causee}}$ (x_2) . . . (x_n)

The effect of this rule is that (i) an extra argument (x_0) with the function of Causer is added to the input predicate-frame, (ii) the first argument of the input predicate-frame gets the function of Causee in the output predicate-frame, and (iii) an extension marker E is added to the input predicate to signal the causative status of the output predicate.

In a great many languages, the extension marker E is an affix. This is for instance the case in Hungarian. Consider the following examples:

(45) (a) *Mari kimos-t-a a ruhák-at.*
Mary wash-past-3s the clothes-acc
'Mary washed the clothes'
(b) *Mari-val kimos-at-t-am a ruhák-at.*
Mary-instr wash-caus-past-is the clothes-acc
'I had Mary was the clothes'

Other examples are- *épít* 'build' — *építtet* 'have build', *sétál* 'walk' — *sétáltat* 'take for a walk', *olvas* 'read' — *olvastat* 'have read', *tart* 'hold' — *tartat* 'have hold', *ül* 'sit' — *ültet* 'have sit down'.

The following predicate formation rule by and large accounts for causative formation in Hungarian:[10]

(46) CAUSATIVE PREDICATE FORMATION IN HUNGARIAN
input: pred_V ([+con] x_1) . . . (x_n)
output: pred_V-E ([+con] $x_0)_{\text{AgCauser}}$ $(x_1)_{\text{Causee}}$ . . . (x_n)
E = *-(t)at/-(t)et*
meaning: 'x_0 brings it about that the state of affairs designated by the input predicate-frame takes place'

Note that this rule accounts, inter alia, for the introduction of the causative formative suffix and the extra argument. It also accounts for there being different controllers of the state of affairs designated by the input and output predicate-frames. Compare (47a) and (47b), where *szándékosan* 'intentionally' depends on the will of Mari in (47a), and on the will of Péter in (47b):

[10] See Hetzron (1976) for a discussion of semantic properties of input and output predicates, constraints on the Causee, and case assignment.

(47) (a) *Mari szándékosan kimos-t-a a ruhák-at.*
Mary intentionally wash-past-3s the clothes-acc
'Mary intentionally washed the clothes'
(b) *Péter szándékosan kimos-at-t-a a ruhák-at Mari-val.*
Peter intentionally wash-caus-past-3s the clothes-acc Mary-instr
'Peter intentionally had Mary wash the clothes'

5.3. Valency reduction

By valency reduction we understand an operation that takes an n-place predicate as input and gives as output the n-place predicate minus one argument.[11] For instance (48) and (49):

(48) FIRST ARGUMENT REDUCTION
input: $\text{pred}_V\ (x_1)\ (x_2) \ldots (x_n)$
output: $\text{pred}_V\text{-R}\ (x_2) \ldots (x_n)$

(49) SECOND ARGUMENT REDUCTION
input: $\text{pred}_V\ (x_1)\ (x_2) \ldots (x_n)$
output: $\text{pred}_V\text{-R}\ (x_1) \ldots (x_n)$

An example of a predicate formation rule that involves the reduction of the first argument of input predicate-frames is the formation of a class of process predicates in Hungarian. Consider the following two examples:

(50) (a) *János zárja az ajtó-t*
John close the door-acc
'John closes the door'
(b) *az ajtó zár-ódik* (**János által*)
the door close-R John by
'the door closes'

Other examples of this pair of predicates are: *rak* 'put' — *rakódik* 'be deposited', *csinál* 'make' — *csinálódik* 'be done', *pácol* 'pickle' — *pácolódik* 'be in the process of pickling', *ír* 'write' — *íródik* 'be written', and *elad* 'sell' — *eladódik* 'be sold'. The following predicate formation rule may account for the relation between these pairs of verbs:

[11] See Mackenzie (1985b) for argument reduction in nominalizations, and Kahrel (1985) for an alternative 'step by step' treatment of derived intransitives.

(51) INTRANSITIVE PREDICATE FORMATION IN HUNGARIAN

input: $pred_V$ ([+con] x_1)$_{Ag}$ ([±tel] x_2)$_{Go}$
output: $pred_V$-R ([−con], [±tel] x_2)$_{Proc}$
R = *-ódik/-ődik*
meaning: 'the predicate $pred_V$ is relevant only to (x_2)'

The formation of verbal reflexives and reciprocals in Hungarian constitutes an example of second argument reduction. Consider the following examples:

(52) (a) *a borbély borotválja Feri-t*
the barber shave Feri-acc
'The barber shaves Feri'
(b) *Feri borotvál-kozik* (**magá-t*)
Feri shave-R himself-acc
'Feri shaves himself'

A predicate formation rule that derives verbal reflexives/reciprocal from transitive predicates in Hungarian may have the following form (cf. Dik 1983):

(53) REFLEXIVE/RECIPROCAL PREDICATE FORMATION IN HUNGARIAN

input: $pred_V$ ([+con] x_1)$_{Ag}$ ([±tel] x_2)$_{Go}$
output: $pred_V$-R ([+con], [±tel] x_1)$_{Ag}$
R = *-kozik/-kezik/-közik*
meaning: 'the relation expressed by $pred_V$ applies to x_1'
condition: input predicate must be a predicate which can take an animate Goal.

Rule (53) may also account for the relation between the following pairs of predicates: *törül* 'dry' — *törülközik* 'dry oneself', *beírat* 'have something listed' — *beíratkozik* 'enrol', *bemutat* 'introduce' — *bemutatkozik* 'introduce oneself', *elígér* 'promise' — *elígérkezik* 'engage oneself', and *ölel* 'embrace' — *ölelkezik* 'embrace each other'.

Apart from operations such as valency reduction, rules (51) and (53) also account for the following difference between the one-place output predicates: the first argument in an output predicate-frame of rule (51) will be a non-controller in a Process state of affairs, whereas the first argument of an output predicate-frame of rule (53) will be the controlle in an Action state of affairs.

The two types of output predicate-frame have in common the property that they can be used in telic states of affairs without any need for a satellite

to set the terminal point. With respect to this property, the output predicate-frames differ from other one-place predicates such as *sétál* 'walk' and *hull* 'fall', in that the latter predicates need a satellite to set the terminal point in a telic state of affairs. Consider the following sentences:

(54) *János megborotválkozott*
John shaved himself'

(55) *az ajtó bezáródott*
'the door closed'

(56) *Mari sétált* (*a pályaudvarra*)
'Mary walked' ('to the station')

(57) *a levél hullott* (*a földre*)
'the leaf fell' ('to the ground')

Thus, the contingent features and their different values together define a semantic cross-classification of one-place predicates in Hungarian as given in the following figure:[12]

	+ con	— con
+ tel	borotválkozik (54)	záródik (55)
— tel	sétál (56)	hull (57)

Fig. 4. A typology of one-place predicates in Hungarian

5.4. Semantic function shift

In the application of the intransitive predicate formation rule in Hungarian (cf. (51)), we have seen that the argument with the semantic function of Goal in the input predicate-frame becomes the argument with the semantic function of Processed in the output predicate-frame. We refer to this shift as "semantic function shift". Many predicate formation rules have several effects on input predicate-frames, for instance change in valency number together with semantic function shift together. The following example illustrates a semantic function shift between the second and third arguments of a predicate. There is no valency extension or valency reduction involved.

Many languages possess oppositions corresponding to the following pair of expressions:[13]

[12] The typology of one-place predicates can be made more precise if the inherent features [dyn] and [mom] are also taken into consideration.

[13] See for instance: Anderson (1971), Comrie (1985), Dik (1980), Hopper & Thompson (1980), and Moravcsik (1978). For a discussion of this type of predicate within FG: Bolkestein (1983, 1985), Dik (1980: ch. 2), De Groot (1984).

(58) (a) *János vaj-at ken a kenyér-re.*
John butter-acc spread the bread-subl
'John spreads butter on the bread'
(b) *János vaj-jal keni a kenyer-et.*
John butter-instr spread the bread-acc
'John spreads the bread with butter'

Unlike similar examples from other languages, construction (58b) in Hungarian does not necessarily have an element of "completeness'. Both constructions, (58a) and (58b), can have an holistic reading. This is the case when perfective aspect has been applied (see De Groot 1984 for further discussion). The difference in the expressions (58a) and (58b) in Hungarian rather is the choice of the entity affected/effected by the predicate. Based on Dik's (1980) proposal for Dutch, the following formation rule may account for the relation between constructions such as (58a) and (58b):

(59) VALENCY REARRANGEMENT PREDICATE FORMATION IN HUNGARIAN
input: $pred_V$ $(x_1)_{Ag}$ $([\pm tel]$ $x_2)_{Go}$ $(x_3)_{Loc/Dir}$
output: $pred_V$ $(x_1)_{Ag}$ $([\pm tel]$ $x_3)_{Go}$ $(x_2)_{Instr}$
meaning: 'the Action expressed by $pred_V$ affects x_3, by using x_2'

Note the shift of the input Goal to Instrument alongside the shift of the input Location/Direction to Goal. Note also that the features [tel] binds different entities in the input and output predicate-frames. This accounts for the application of different preverbs and different holistic interpretations with perfective aspect. Consider:

(60) (a) *János rá-keni a vaj-at a kenyér-re.*
John pf-spread the butter-acc the bread-subl
'John will spread (all) the butter on the bread'
(b) *János a vaj-jal meg-keni a kenyer-et.*
John the butter-instr pf-spread the bread-acc
'John will spread (all the slices of) the bread with the butter'

We may therefore conclude that an account of the relation between the Hungarian constructions under discussion can best be given in terms of telicity binding different entities.

5.5. Categorial change of input predicate

In languages across the world the following triplet of expressions often occurs:

(61) (a) *be happy* (State)
(b) *get happy* (Process)
(c) *make happy* (Action)

In most cases the adjective can be considered the basic form from which the inchoative/pseudo passive and transitive/causative form can be derived. The rules for doing this can in general be formulated in the following way:[14]

(62) input: $pred_A\ (x_1)_\emptyset$
output 1: $pred_V\text{-}D\ (x_1)_{Proc}$
output 2: $pred_V\text{-}D\ (x_2)_{Ag}\ (x_1)_{Go}$

This pattern is also found in Hungarian. Consider the following sentences:

(63) (a) *Mari szép.*
Mary pretty
'Mary is pretty'
(b) *Mari szép-ül-ø.*
Mary pretty-inch-3s
'Mary is getting prettier'
(c) *Mari szép-ít-i magá-t.*
Mary pretty-caus-3s herself-acc
'Mary beautifies herself'

Examples of other such triplets are: *rövid* 'short' — *rövidül* 'shorten' — *rövidít* 'shorten', *szabad* 'free' —*szabadul* 'be freed' — *szabadít* 'liberate', *vak* 'blind' — *vakul* 'go blind' — *vakít* 'put somebody's eyes out', *mély* 'deep' — *mélyül* 'deepen' — *mélyít* 'deepen', *kék* 'blue' — *kékül* 'become blue' — *kékít* 'make blue'. The following predicate formation rules may account for the relation between these predicates:

[14] The relation between the three constructions does not necessarily have to be the one presented here. There may also be an ordering in the formation rules, for instance:
I. State → Process
II. Process → Action

(64) DEADJECTIVAL PREDICATE FORMATION IN HUNGARIAN

input: [−dyn] $pred_A$ ([−con] x_1)ø

output 1: [+dyn] pred-D_V ([−con], [±tel] x_1)$_{Proc}$
D = *ul* / *ül*

meaning: 'the property expressed by $pred_A$ is presented as coming about through a process'

output 2: [+dyn] pred-D_V ([+con] x_2)$_{Ag}$ ([±tel] x_1)$_{Go}$
D = *-it*

meaning: 'x_2 brings it about that the property expressed by $pred_A$ applies to x_1'

The effect of this rule is that (i) the categorial status of the output predicates is Verb, (ii) the inherent feature of those predicates is [+dyn], (iii) one of the arguments is bound by the feature [tel], and (iv) the marker -D is added to the input predicate to signal the deadjectival status of the output predicate.

6. Conclusions

In this paper it has been shown that Functional Grammar offers a framework that can account for many morphological and semantic aspects of predicate formation.

(i) Distinctions between morphological categories can be considered to be theory-internal. We may summarize this in the following way: "roots/stems" are listed in the lexicon, derivational morphemes in the predicate formation component, and inflectional morphemes in the expression rules component.

(ii) Given the formalism of predicate-frames as defined in FG (category, valency, semantic functions of the arguments, and specifications for the sets of states of affairs they designate), the model can account for a number of predicate formation rules in terms of:

(a) change of category
(b) change in valency
(c) semantic function shift
(d) different distribution of features

(iii) Predicate formation as described above does not require an extension of the theory: predicate formation rules make use of an available formalism. The rules proposed for Hungarian are not 'ad hoc' rules, because basically the same rules can be formulated for many other languages. Moreover, the model can account for several aspects of different types of predicate formation within a language in a unifying way. As an example I mention the role semantic function shift and the assignment of the feature [tel] play in the derivation of one-place and three-place predicates.

References

Anderson, Stephen R.: On the role of deep structure in semantic interpretation. In: Foundations of Language 7 (1971). 378–396.

Anderson, S. R.: Where is morphology? Linguistic Inquiry 13 (1982), 571–612.

Bolkestein, A. M.: The role of discourse in syntax: evidence from the Latin Nominativus cum Infinitivo. In: Ehlich K. and van Riemsdijk, H. C. (eds): Connectedness in Sentence, Discourse and Text. Tilburg: Dept of Language an Literature, Tilburg University 1983, 111–140.

Bolkestein, A. M.: Discourse and Case-Marking: Three-place predicates in Latin. In: Touratier, Chr. (ed.): Syntax et Latin. Université de Province, Aix en Provence 1985, 191–225.

Bolkestein, A. M., Combé, H. A., Dik, S. C. Groot, C. de, Jadranka Gvozdanovic, Rijksbaron, A. and Vet, Co: Predication and Expression in Functional Grammar. London–New York, Academic Press 1981.

Bolkestein, A. M., Groot, de C. and Mackenzie, J. L. (eds): Syntax and Pragmatics in Functional Grammar. Functional Grammar Series 1. Foris, Dordrecht 1985a.

Bolkestein, A. M., Groot, C. de and Mackenzie, J. L.: Predicates and Terms in Functional Grammar Series 2. Dordrecht: Foris, Dordrecht 1985b.

Brigden, N.: Towards a functional grammar of aspect in Finnish. In: Groot and Tommola (eds), 179–198.

Bybee, J. L.: Morphology. Typological Studies in Language 9. Benjamins, Amsterdam 1985.

Comrie, B.: Aspect. Cambridge University Press, Cambridge 1976.

Comrie, B.: Causative verb formation and other verb-deriving morphology. In: Shopen (ed.), 309–348.

Dik, S. C.: Functional Grammar. North-Holland Linguistic Series 37. North-Holland, Amsterdam 1978 (Third printing: Foris, Dordrecht 1981).

Dik, S. C.: Funktionele Morfologie. In: Hoekstra T. and van der Hulst (eds): Morfologie in Nederland. Glot-special, Leiden 1979, 72–100.

Dik, S. C.: Studies in Functional Grammar. Academic Press, London–New York: 1980.

Dik, S. C.: Over de behandeling van niet-produktieve regelmatigheden. In: Forum der Letteren 22 (1981), 29–50.

Dik, S. C.: On the status of verbal reflexives. In: Communication and Cognition 16 (1983), 39–63. (= Tasmowski, L. and Willems, D. (eds): Problems in Syntax. Studies in Language 2. Gent: Plenum/Communication and Cognition, Gent 1983, 321–255.)

Dik, S. C.: Formal and semantic adjustment of derived constructions. In: Bolkestein, Groot and Mackenzie (eds) 1985a, 1–28.

Dik. S. C.: Valentie en valentie-operaties in Funktionele Grammatika. In: Dik (ed.): 1985b, 95–114.

Dik, S. C.: The theory of Functional Grammar. Part I: The Simple Predication. In preparation.

Dik, S. C. (ed.): Advances in Functional Grammar. Foris, Dordrecht 1983.

Dik, S. C. (ed.): Valentie in Funktionele Grammatika. In: TTT (Tijdschrift voor Taal- en Tekstwetenschap) 5 (1985), 2.

Groot, C. de: The structure of predicates and verb agreement in Hungarian. In: Daalder S. and Gerritsen, M. (eds.) Linguistics in the Netherlands North-Holland, Amsterdam 1981, 149–158.

Groot, C. de: Typology of states of affairs. In: Bennis, H. and Van Lessen Kloeke, W. U. S. (eds): Linguistics in the Netherlands. Foris, Dordrecht 1983, 73–81.

Groot, C. de: Totally affected. Aspect and three-place predicates in Hungarian. In: Groot, C. de and Tommola, 4. (eds), 1984, 133–151.

Groot, C. de: Predicates and features. In: Bolkestein, Groot and Mackenzie (eds), 1985b, 71–84.

Groot, C. de and Tommola, H. (eds): Aspect Bound. A Voyage into the Realm of Germanic, Slavonic and Finno-Ugrian Aspectology. Foris, Dordrecht 1984.

Hetzron, R.: On the Hungarian causative verb and its syntax. In: Shibatani, M. (ed.): Syntax and Semantics 6: The Grammar of Causative Construction. Academic Press, London–New York 1976, 371–398.

Hoekstra, T. Teun, Hulst, H. van der and Moortgat, M. (eds): Perspectives on Functional Grammar. Foris, Dordrecht (Glot 3 (1980), 3–4).

Hopper, P. and Thompson, S. A.: Transitivity in grammar and discourse. In: Language 56 (1980), 251—299.
Junger, Judith: Morphological causatives in Modern Hebrew. In: Bolkestein, Groot and Mackenzie (eds), 1985, 235—257.
Junger, J.: Valentiereduktie in het Modern Hebreews. In: Dik (ed.) 1985, 141—160.
Kahrel, P.: Some aspects of derived intransitivity. In: Working Papers in Functional Grammar 4 (1985).
Mackenzie, J. L.: Nominalization and valency reduction. In: Bolkestein, Groot and Mackenzie (eds), 1985, 29—47.
Mackenzie, J. L.: Genominaliseer. In Dik (ed.), 1985, 177—198.
Moravcsik, Edith A.: On case-marking of objects. In: Greenberg, J. H. (ed.): Universals of Human Language. Vol. 4. Stanford University Press, Stanford 1978.
Samuelsdorff, P.: O.: The treatment of morphology in a Functional Grammar with special reference to Swahili. In: Folia Linguistica 16 (1982), 385—398.
Schaaik, G. van: A functional analysis of aspects of Turkish grammar. MA Thesis, Institute for General Linguistics, University of Amsterdam 1983.
Schaaik, G.: Valentiereduktie in het Turks. In: Dik ed. 1985, 127—139.
Shopen, T. (ed.): Language Typology and Syntactic Description. Part III. Grammatical Categories and the Lexicon. Cambridge University Press, Cambridge 1985.
Vester, Elseline: Instrument and Manner Expressions in Latin. Van Gorcum, Assen 1983.
Vet, C.: Temps, aspect et adverbes de tempes en francais contemporain. Droz, Geneve 1980.
Vet, C.: Passive, reflexive and causative predicate formation in French. In: Bolkestein, Groot and Mackenzie (eds), 1985, 49—69.
Watters, J. R.: The place of morphology in FG: the case of Ejagham verb system. In: Bolkestein, Groot and Mackenzie (eds), 1985, 85—104.
Williams, E.: Argument structure and morphology. In: Linguistic Review 1 (1981), 81—114.
Workgroup on Functional Grammar: On the Functional Grammar of Teaching Verbs. In: Hoekstra, Hulst and Moortgat (eds), 1981, 203—231.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 93–107 (1986)

PROBLEMS IN THE MORPHOLOGICAL ANALYSIS OF COMPLEX LEXICAL ITEMS

D. KASTOVSKY

(University Wien, Austria)

1.1. The study of word-formation in the past thirty years reflects the influence of two different, but in a way complementary linguistic traditions. European contributions to the field, and in particular those of Marchand (e.g. Marchand 1969, 1974) and his pupils, are characterized by a Saussurean, sign-oriented bias, where functional and semantic considerations outrank other criteria. Morphemes are regarded as minimal linguistic signs "based on a significant/significate [...] relationship" (Marchand 1969, 1) and are consistently distinguished from "free forms" or "pseudo-morphemes" such as *make* and *out* in *make out*, which have no sign character. Word-formation is restricted to the description of lexemic sign-combinations because "only meaningful units lead to analogic new formations" (Marchand 1969, 2), and the notion of "motivation" is therefore central to this approach (cf. also Kastovsky 1987).

The majority of the contributions coming from the U. S., on the other hand — and this is also true of the more recent generative treatments coming from Europe —, emphasize formal, i.e. phonological and distributional aspects, most often with a concomitant neglect of semantic questions. We might therefore even speak of a "(Neo-) Bloomfieldian" bias, not in the sense of "taxonomic", as this term was used by the first generation of generative linguists, but with the implication that the dominance of formal, asemantic criteria, characteristic of (Neo-)Bloomfieldian linguistics, is also a typical feature of these more recent generative approaches. Morphological analyses are usually based on phonological and/or distributional criteria, regardless of whether the resulting constituents have sign character in the above sense or not. As a consequence, morphemes or "formatives" are not defined as form/meaning correlations, but rather as those minimal entities that enter into the statement of rules at the phonological and/or syntactic level, regardless of whether these entities are meaningful or not. The following examples will, I hope, corroborate this assessment.

1.2. In Chomsky—Halle (1968, 94), it is purely phonological considerations such as the simplification of certain phonological rules (e.g. Main Stress

Rule, Alternating Stress Rule) that motivate the introduction of the =-boundary, and with it morphological segmentations such as

(1) *re=ceive, de=ceive, re=sist, con=sist, com=pre=hend* etc.

And in Halle (1973, 10) we find segmentations such as

(2) *serendip+i+ty, vac+ant, tot+al, bro+ther, hand+some, be+lieve, trans+form+at+ion+al*

which are, incidentally, not justified by the author, but which, if they could be justified at all, could only be validated by purely distributional criteria. In both examples, most of the resulting constituents do not have sign character. This type of analysis, and even quite a few of the examples themselves, have their counterparts in "descriptive analysis" as introduced by Bloomfield and practised in the 1940s and 1950s (cf. Bloomfield 1933, 209, 241; Harris) 1953, 161; Nida (1949, 162, 191). Not surprisingly, this approach had been criticized for its neglect of the semantic aspect by European structuralists, e.g. Marchand (1951, 94) or Frei (1954, 139, 141). There is no denying that segmentations like those discussed above may be helpful in accounting for the phonological — and occasionally even morphological — behaviour of the lexical items in question. But so far the usual wording of the analyses referred to above suggests that such patterns are productive — which they are not — and that they also involve semantic compositionality, which again is not the case. Consequently, if for phonological reasons analyses such as (1) or (2) are maintained, they must be consistently kept apart from word-formation proper by assigning idiom status to the respective items. But so far, with the exception of Pesetsky (1985, 210), who postulates "rules of idiosyncratic interpretation" for such cases, they have been treated as being morphologically on the same level as

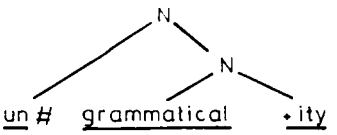
(3) *re # write, de # louse, trans # alpine,* etc.

1.3. Another example illustrating the dominance of phonology over morphology is provided by the Level Ordering Hypothesis as developed by Siegel (1974) and Allen (1978) and the bracketing paradox it causes, cf. Pesetsky (1985, 202 f.).

It has been suggested that on the basis of their phonological behaviour two classes of affixes should be distinguished, viz. # -affixes and +-affixes. These are attributed to two different distributional levels, which are ordered such that no +-affix may be attached to an item to which a # -affix has already been added. This means that no # -affix occurs within the scope of

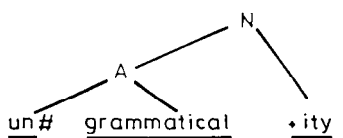
a +-affix. Now given that *un-* is a # -affix and *-ity* a +-affix, the only morphological bracketing possible is the following:

(4)



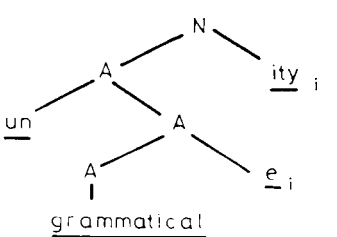
But, as Pesetsky (1985, 202) quite correctly points out, this analysis is contradicted by the derivational facts — *un-* is generally not added to nouns, only to adjectives —, and the semantics of the combination — the meaning is 'state, fact of being ungrammatical'. These facts would require the bracketing which, however, is ruled out by the Level Ordering Hypothesis.

(5)



It is bracketing paradoxes like these that lead Pesetsky to postulating at least two levels of representation for word-formation as well as for syntax: Surface Structure as input to phonology, and Logical Form as input to semantic interpretation. (4) would then represent the bracketing at the level of (morphological) Surface Structure, (5) at the level of Logical Form. The two levels are related by a mapping mechanism which is equivalent to a generalized version of the Quantifier Rule as postulated by May (1977), and which in principle changes the dominance relation within the tree structure by raising the suffix *-ity* out of the scope of *un-* leaving behind a trace (e_i), i.e. a more elaborate version of (5) would be (6) in Pesetsky's framework:

(6)



The assumption of two levels of representation for word-formation is of course not completely new. It was part of the theory of transformationalist word-formation, where a syntactic/semantic representation was converted into a morphological surface structure by syntactic or pre-lexical transforma-

tions.[1] And we also find it in Marchand's work, which is not strictly generative-transformational, as a distinction between "morphological surface structure" and "syntactic deep structure", also with concomitant bracketing differences, cf. Marchand (1965, 57, n. 1) with reference to synthetic compounds like *housekeeping*, and Marchand (1969, 134 ff.) with regard to prefix formations such as *defrost, unsaddle, untie*. There is no explicit mapping mechanism relating these two levels, however; Marchand only speaks of a reinterpretation at the morphological level without providing an appropriate mechanism for this process. Within lexicalist word-formation, however, Pesetsky's two-level approach thus is, as far as I can see, an innovation. And it seems to me that within the lexicalist framework it may well help to overcome certain shortcomings inherent in recent analyses of prefixations, which are absent in the less formalized analysis of Marchand, and which will be discussed in the following. The major question to be asked in this connection will be: are prefixes really heads?

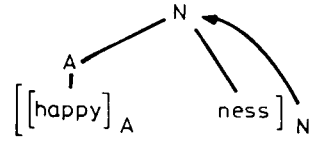
2.1. I will begin this section by showing that two rather recent, widely accepted innovations in lexicalist word-formation — Williams' (1981) theory of (right-hand) headedness and Lieber's (1981) feature percolation conventions — together with the handling of conversions in the given framework and a certain mechanistic bias are responsible for treating prefixes as heads in certain types of combinations. Such an analysis however, has certain unwanted consequences; in particular, one and the same prefix may act as head and non-head in certain, intuitively closely related formations, which clearly leads to internally contradictory results. I will then compare this analysis with the semantically oriented approach advocated by Marchand and will finally make some suggestions as to how the Lieber—Williams analysis might be improved to be intuitively more convincing.

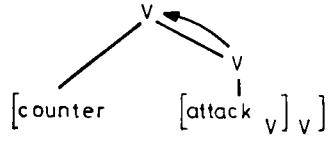
2.2. In Lieber (1981) it is postulated that word-structures start out as binary unlabelled trees, into which, governed by appropriate selection restrictions, terminal lexical items — i.e. stems, roots and affixes — are inserted. A set of 4 feature percolation conventions provides labels for the nodes in these abstract trees (Lieber 1981, 47 ff., 1983, 252 ff.). Of special interest in this connection are conventions II and III. According to these, all affixes with the exception of those belonging to the so-called "null-category class", act as heads of the respective combination and percolate their features to the next highest branching node. This is true not only of suffixes, but also of prefixes, cf.:[2]

[1] Cf. the role of Predicate Raising, NP-Adjunction, NP-Copying in Kastovsky (1982, 234ff.)

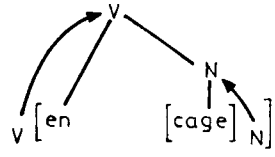
[2] The representations in (7) are based on Lieber (1983)

(7) (a) Convention II[3] (suffix=head)





(b) Convention III[4] (prefix="null-category class"; stem=head)



(c) Convention II (prefix=head)

Lieber's Convention II had been formulated in direct response to Williams' Right Hand Head Rule:

> In morphology, we define the head of a morphologically complex word to be the right-hand member of that word (Williams 1981, 248).

As already noted by Williams himself (1981, 249), there appeared to be certain exceptions to this rule, e.g. the prefix *en-*, which "systematically converts nouns and adjectives into verbs, thus displaying the behaviour of a head: *enrage, endear, ennoble, encase*". Further examples from German and English led Lieber to conclude that these were not exceptions at all and to stipulate that the position of the head was affix-specific, i.e. that there was no universal generalization as postulated by Williams. The result was Convention II according to which prefixes are heads when they are characterized by a category feature.

[3] Convention II: "All features of an affix morpheme, including category features, percolate to the first branching node dominating that morpheme" (Lieber 1983, 253)

[4] Convention III: "If a branching node fails to obtain features by Convention II, features from the next lowest labeled node automatically percolate up to the unlabeled branching node"

2.3. Williams' Right Hand Head Rule as a universal is clearly inadequate, and Lieber's reaction to it is justified as such, although it is by no means unproblematic either, as we shall see below. The head/modifier order in morphological composites is a matter of language type. Thus, while in English compounds the right hand constituent usually acts as head, this is not true of French compounds such as

(8) *timbre-poste, ouvre-bouteille, maison de campagne.*

And there are obviously also languages where the normal sequence is head/modifier even with affixes, e.g. in Vietnamese, according to Selkirk (1982, 21). On the other hand, many languages apparently have a general p r e f e r r e d order from a typological point of view, just as in syntax, which may perhaps not always be carried through completely for historical reasons (e.g. due to borrowing), but which nevertheless is the unmarked, normal choice. Thus Marchand claims that "in the system of languages to which English belongs the determinant generally precedes the determinatum. The types which do not conform to this principle are either syntactical compounds (e.g. *father-in-law*) or loan compounds, e.g. *McDonald, Fitzgerald*" (Marchand 1969, 11). This observation indeed holds true for all Germanic languages, and without really conclusive evidence we should be very careful to assume the opposite order — even if it seems to be the only option in the framework used, which brings me back to the analysis of *enrage, endear*, etc. as suggested by Williams and Lieber.[5]

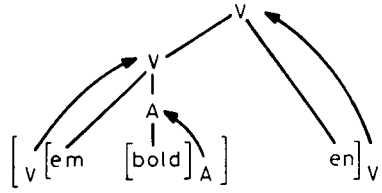
2.4. One crucial observation is that the alleged headhood of a prefix like *en-* leads to an internally contradictory analysis. Lieber and Williams have only considered cases where there is no overt suffixal element that could act as head and thus percolate its features to the appropriate dominating node. But besides such formations, there are formations with a prefix and a suffix, which follow exactly the same semantic pattern as the ones without the explicit suffix, e.g.

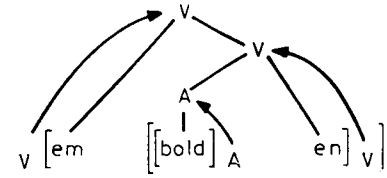
(9) (a) *en-noble* : *em-bold-en* 'to make adj.'
de-louse : *de-sulphur-ate* 'to remove N'
de-gas : *de-gas-ify* 'to remove N'
(b) *ver-riegel-n* : *ver-barrikad-ier-en* 'to block with N'
be-eid-en : *be-eid-ig-en* 'to confirm by N'
be-dach-en : *be-teil-ig-en* 'to provide with N'

[5] It should be pointed out that headhood for prefixes was also questioned for Dutch in a recent article by Trommelen—Zonneveld (1986), although for somewhat different reasons than the ones discussed in the following.

In the left-hand column, the prefix acts as head according to Lieber's and Williams' analysis; but what about the right-hand column? Is the prefix or the suffix the head? Both options seem plausible, cf.:

(10) (a) Suffix as ultimate head





(b) Prefix as ultimate head

In Williams' framework, where prefixes are heads only exceptionally, (10) (a) would definitely have to be chosen as the appropriate analysis, in which case the prefix *en-* acts as head in *ennoble* 'make noble', but as non-head in *embolden* 'make bold', certainly an inconsistency.[6] In Lieber's framework, on the other hand, no such preference seems to exist, so that (10) (b) might be chosen over (10) (a), in which case a prefix like *em-* would always act as head. Also note that one of Williams' arguments in favour of treating *em-* as head was affix potentiation, i.e. the fact that the prefix *en-* feeds suffixation by *-ment*. Due to his Atom Condition, which makes internal structure irrelevant for further derivation, it is the features of the head that govern any further derivation. The choice of *-ment* as nominalizing suffix with *en-*verbs is thus explained, if *en-* is interpreted as head. Examples like *enlightenment*, *?emboldenment*, etc. would seem to corroborate this and thus support solution (10) (b) for Lieber's framework. But unfortunately the data are not that straightforward, since there are also cases such as

(11) *embark* : *embarkation*
encapsulate : *encapsulation*
encarnalize : *encarnalization*
enclose : *enclosure*

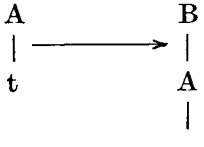
[6] And the existence of a verb *to embold* (cf. Shorter Oxford English Dictionary s. v. *embolden*) increases this problem, despite the latter's being archaic.

It is true that most pure, i.e. unsuffixed *en*-verbs select the nominalizing suffix *-ment*, but with suffixal *en*-verbs it seems to be the suffix which determines the form of the nominalization suffix. This is particularly clear with verbs in *-ate*, *-ize*, *-ify*, which only permit *-ation*. Consequently, if there is such a thing as the Atom Condition, structure (10) (a) — the suffix as ultimate head — will have to be chosen over (10) (b). We are thus back to the dilemma mentioned above. Moreover, such a solution faces problems in the semantic interpretation, which are not discussed, incidentally, by either Lieber or Williams, since both are only interested in the formal derivational mechanism. We will have to postulate two types of semantic rules for *en-*, one where it acts as head, and one where it acts as modifier, although the resulting semantic structures are the same. This is certainly an unwelcome consequence, but one obviously unavoidable in the present framework, except if we find a way to treat prefixes as non-heads in cases such as *enrage*, *ennoble*, etc. And this brings me to the treatment of conversions.

2.5. In English (and in German), there is a word-formation pattern, illustrated by

(12) *clean*$_A$: *clean*$_V$ 'make clean'
gas$_N$: *gas*$_V$ 'convert into/treat with gas'
bag$_N$: *bag*$_V$ 'put into a bag'
gut$_N$: *gut*$_V$ 'remove the guts'

which is commonly called conversion in the traditional handbooks, and which is analyzed as zero derivation by e.g. Marchand. Now it seems to me that the derivative process producing verbs such as *ennoble*, *encage*, *debug*, *unsaddle*, etc. is closely related to conversion, and not only semantically. This fact was overlooked by Lieber and Williams because of the way they treat conversions. Lieber regards conversion not as a derivational process at all but treats pairs like those in (12) as independent lexical items related by a redundancy rule in the lexicon; only the rule of semantic interpretation may be directional. And Williams postulates a special type of derivational process, so-called headless rules, which do not result in a branching structure, a solution also found in Strauss (1982, 54), who argued that "zero morphemes are entirely unnecessary for describing zero derivation . . . zero derivation is simply word-formation without terminal adjunction — zero adds a non-terminal and no terminal":

(13)


Conversions in this framework are thus regarded as a mere change of category, not as a process that is exactly parallel to explicit suffixal derivation. And it is this decision that forces the analysis of *encage*, etc. as containing a prefixal head, instead of treating it as morphologically parallel to cases such as *encapsulate* etc. The reason for this decision seems to be the exclusive reliance on the formal aspect, with a concomitant neglect of semantic considerations. Let us therefore now look at the description of prefixations in a more semantically oriented framework such as Marchand's.

3.1. Central to Marchand's theory of word-formation is the syntagma concept, i.e. the assumption that all word-formations typically consist of a binary determinant/determinatum structure, cf.

(14)

determinant	/	determinatum
steam	/	boat
letter	/	writer
ex	/	president
re	/	write
writ	/	er
steam	/	er

This holds regardless of whether this structure is realized overtly or not, i.e. it is also applied to conversions on the basis of proportional equations such as

(15) (a) *write*: *writ-er* = $cheat_V$: $cheat_V/\emptyset_N$ = 'someone who V-s'
(b) *atom*: *atom-ize* = $cash_N$: $cash_N/\emptyset_V$ = 'to convert into N'
(c) *soft*: *soft-en* = $clean_A$: $clean_A/\emptyset_V$ = 'to make clean'
(d) *pre-atom/ic* (period) = *pre-war*/Ø (*period*) = 'N before N'

Conversions are therefore interpreted as zero-derivations, where Ø functions as a purely formal indicator of a semantically implied, but formally unexpressed element, as in mathematics, which provides the necessary structure for the semantic interpretation. Incidentally, there is often evidence that in an historically earlier period these formations had been characterized by an explicit suffix, which, however, was lost at a later stage.[7] Conversions thus

[7] Thus, the Germanic weak verbs, historically the antecedents of present-day conversion verbs, had been characterized by different stem-formatives, which originally must have had the function of derivational suffixes, but which later became mere indicators of inflectional class, until they finally merged with either the verbal stem or the inflectional endings — i.e. were lost. The same seems to be true of nominal stem-formatives, e.g. *-a* in *hunta* 'hunter', which in OE has to be analyzed as a case/number exponent, and no longer as a stem formative. Zero is thus nothing but a device to indicate the lack of an explicit derivative element, at the same time maintaining the semantic/functional parallelism with explicit suffixation

are not treated as mere categorial shifts, but as regular derivational, suffix-like processes with the same semantic patterning as suffixations.

3.2. The field of word-formation is subdivided into two major subtypes, expansions and derivations. The former satisfy the formula AB=B, i.e. the combination as a whole can be replaced by the determinatum (head), which is an actual or possible lexical item; this is not possible with derivations, where the determinatum is a bound morpheme or zero, cf.

(16) (a) *steamboat* → *boat*: AB = B = Exp.
madman → *man*: AB = B = Exp.
ex-husband → *husband*: AB = B = Exp.
re-write → *write*: AB = B = Exp

(b) *writer* → ?: AB # B = Der.
whiten → ?: AB # B = Der.
redskin ↛ *skin*: AB # B = Der.
(redskin/Ø)
pickpocket ↛ *pocket*: AB # B = Der.
(*pickpocket*/Ø)

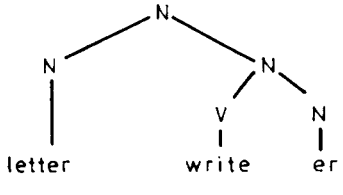
Since *redskin* and *pickpocket* do not satisfy the formula AB = B, they are treated as (zero-)derivatives or pseudo-compounds, in order to distinguish them from regular compounds such as *steamboat*.

3.3. Compounds are further subdivided into

a) genuine compounds (the determinatum is a simple lexical item), e.g. *steamboat*; *drawbridge*;
b) synthetic compounds (the morphological determinatum is a deverbal derivative), e.g. *letterwriter*, *chimney-sweep*, *housewarming*;
c) pseudo-compounds (the superficial structure is that of a), i.e. of a regular compound, but AB = B does not apply), e.g. *pickpocket*, *hunchback*, *redskin*, *stagemanage* (< *stagemanager*), *proofread* (< *proofreading*).

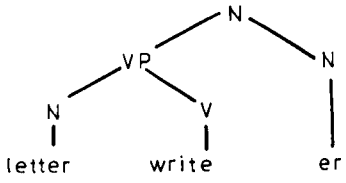
Synthetic compounds are characterized by two types of bracketing, cf. the discussion of *ungrammaticality* in **1.3.** above. From a purely morphological point of view, we get the structure (17),

(17)



with *writer* acting as head of the whole combination. Semantically speaking, however, the compound is analyzed as 'someone who writes (a) letter(s)', which is reflected by the bracketing

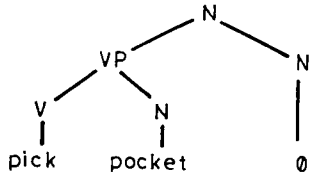
(18)



Here the suffix *-er* acts as the head of the whole combination, which thus makes it look like a derivative from a phrase. As already mentiond, Marchand provides no mechanism to relate these two bracketing structures; this has been done by Pesetsky (1985, 233 ff.), however, among others also for synthetic compounds which he calls "deverbal compounds". Pesetsky only discusses action nouns like the by now almost proverbial *pasta eating*, but similar provisions will have to be made for synthetic agent nouns and other types of synthetic compounds.

Pseudo-compounds, on the other hand, only allow one type of bracketing in Marchand's analysis, viz. (19),

(19)



which clearly identifies them as derivatives at all levels.

3.4. Prefixations can also be subclassified into genuine prefixations, synthetic prefixations and pseudo-prefixations (Marchand 1969, 137); moreover, prefixes may represent three different functions: adjectival (e.g. *ex-husband* 'former husband'), adverbial (e.g. *rewrite* 'write again'), and prepositional (e.g. *pre-atomic age* 'age before the atomic age', *encage* 'put into (*en-*) a cage').

(20) (a) genuine prefixations: *rewrite, ex-husband, hyper-accurate* (*rewrite* → *write*)
(b) synthetic prefixations: *re/wire-∅, under/mine-∅* (*rewire* → *wire/∅*)
(c) pseudo-prefixations: *encage, encarnalize, disbar, outbid, untie, demilitarize, defrost* (*encage* ↛ *cage, untie* ↛ *tie*)

In synthetic prefixations, the determinatum is itself derived, but it need not necessarily exist as a lexicalized unit; thus, an analysis as expansion on the basis AB = B is in principle always possible, cf.

(21) (a) morphological analysis
$[re\ [[wire]_N\ \emptyset_V]_V]_V$ 'wire again'
(b) semantic analysis
$[[re\ [wire]_N\ ]_{NP}\emptyset_V]_V$ 'provide with new wires'

It should be added, however, that synthetic prefixations of this kind seem to be rather rare compared to synthetic compounds.

Pseudo-prefixations, finally, are not analyzable on the basis of AB = B, i.e. *to untie* is, semantically speaking, not a hyponym of *tie*, but rather means 'undo the tiedness'; *to encage* is not a subcategory of *cage* sb., but means 'put into a cage'. These are thus all basically derivatives, either characterized by an explicit suffix or by zero, with the latter being the more frequent option.

Pseudo-prefixations can further be subdivided on the basis of their internal bracketing:

(22) (a) en-cage/Ø un-saddle/Ø en-capsul/ate

pre-atom/ic (age) post-war/Ø (period)

i.e. $\left[[\text{prep} + \text{N}]_{pp}\left\{\begin{matrix}\emptyset_V\\ \text{aff}\end{matrix}\right\}\right]_V \left[[\text{prep} + \text{N}]_{pp}\left\{\begin{matrix}\emptyset_A\\ \text{aff}_A\end{matrix}\right\}\right]_A$

(b) de-frost-Ø un-tie-Ø dis-arm-Ø de-militar-ize

In (22) (a), the basis is a prepositional group followed by an explicit or implicit derivational element, i.e. we get a clearly hierarchical constituent structure. In (22) (b), on the other hand, prefix and noun or adjective do not form a construction, but prefix and suffix together express the notion of reversativity or privativity. They thus are a kind of discontinuous affix, or "circumfix", as Hansen (1980) has called it.

Marchand now argues that structures such as (22) (b), arrived at on the basis of a semantic analysis, and thus representing or at least reflecting the underlying semantic structure, do not conform to the typical morphological prefixal patterns of English; there is, therefore, "a strong tendency in lan-

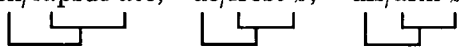
guage to adjust morphological syntagmas in opposition to their syntactic structure as revealed by content analysis" (Marchand 1969, 136). Consequently,

> at the morphological level, the verbs *disarm* and *defrost* come to be treated as prefixal combinations with primary verbs so that *disarm* is parallel to *disjoin*, *defrost* parallel to *desegregate*. The deverbal prefixations are themselves formed as if based on the semantic type 'connect, segregate in reverse' (Marchand 1969, 136).

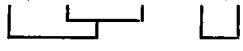
The same type of reanalysis is also suggested for the derivations in (22) (a), so that we arrive at a morphological surface structure (23), which, in fact, is that of synthetic prefixations with the exception of deverbal verbs like *untie*, which are structurally like *rewrite*.

(23) en/cage-Ø, un-saddle-Ø,

en/capsul-ate, de/frost-Ø, dis/arm-Ø,



de/militar-ize, un/tie



4.1. Marchand's analysis as it stands can of course not be incorporated into the lexicalist framework developed by Lieber, Williams, and others, nor can it simply replace it, because it lacks the necessary formal apparatus. But it contains a number of observations that suggest a possible solution to the problem discussed in the second section above.

The first crucial observation is that at the (morphological) surface level (Pesetsky's S-Structure), the bracketing is always $[\text{Prefix} + X]_X$, where X may be simple or complex, but always acts as head, i.e. there is no need to invoke headhood for prefixes. There is thus a difference between pseudo-compounds and pseudo-prefixations, which in Marchand's analysis is not quite obvious. Pseudo-compounds are derivations also surface-structurally, cf. (19), while pseudo-prefixations, at least if one accepts Marchand's theory of morphological reinterpretation, are prefixations (= expansions), and mainly synthetic ones, from a morphological point of view, cf. (23).

The second point is that the morphological surface bracketing of prefixations does not necessarily provide the information necessary for their correct semantic interpretation, so that some restructuring is required, cf. (4), (5) and (6) in **1.3.**

Thirdly and perhaps most importantly, (23) indicates that the second constituents in formations such as *encage* should not be interpreted as nouns,

but rather as — existing or possible — conversion verbs, so that the prefix can act as determinant and need not be interpreted as head causing a change in word class affiliation. Whether conversion is interpreted as zero derivation or along the lines of Strauss or Lieber, is a secondary matter, although to me it seems more plausible and consistent to integrate "conversion" somehow into the binary branching system and the percolation conventions suggested by Lieber.

To interpret cases such as *encage, defrost, unsaddle* as — surface-structurally speaking — containing a conversion verb as head and a prefix as determinant is corroborated by the semantic/functional parallelism with other types of formations, in particular those already mentioned by Marchand, viz. *dis/arm-Ø* : *dis/join, en/cage-Ø* : *en/wrap.* But there are further parallels, e.g. with post-particle (i.e. phrasal) verbs, where the particle has the same determining function as the prefix, cf. *en/cage-Ø* : *cage-Ø/in*; and there is of course the semantic parallelism between simple conversion and prefixal derivatives in all relevant patterns, cf. *to skin/Ø* : *to defrost* 'remove N', *to bottle/Ø* : *to encage* 'to put into N', etc.

4.2. If conversion is regarded as a regular derivative process that is integrated into the feature percolating system developed by Lieber and now widely accepted, and if prefixal derivatives of the type *encage, defrost* etc. are analyzed as containing conversion verbs as their second members, then there is no need to treat the prefixes as heads. As a result, we obtain a more consistent and homogeneous description. On the other hand, such a description will have to be complemented by some kind of restructuring device as developed in Pesetsky (1985), in order to derive the appropriate bracketing structure that can serve as input to the semantic interpretation. In other words, there will be a bracketing difference between surface structure and deep structure (Logical Form), but not one where the prefix will be a head at one level and a modifier at the other; prefixes are not heads.

References

Allen, M.: Morphological Investigations. Ph. dissertation, University of Connecticut, 1978.
Bloomfield, L.: Language. New York 1933.
Chomsky, N. and Halle, M. The Sound Pattern of English. New York 1968.
Frei, H.: Critères de délimitation. In: Word 10 (1954), 136—145.
Halle, M.: Prolegomena to a theory of word-formation. In: Linguistic Inquiry 4 (1973), 3—16.
Hansen, K.: Probleme der Konstituentenanalyse von Wortbildungen im Englischen. In: Hansen, K. and Neubert, A. (eds.), 12—26.
Hansen, K. and Neubert, A. (eds.): Studien zur Lexik und Grammatik der englischen Sprache der Gegenwart. Martin Lehnert zum 70. Geburtstag. Akademie der Wissenschaften der DDR. Zentralinstitut für Sprachwissenschaft. Linguistische Studien. Reihe A, Arbeitsberichte 67. Berlin 1980.

Harris, Z. S.: Methods in Structural Linguistics. Chicago 1953².
Kastovsky, D.: Wortbildung und Semantik. Düsseldorf 1982.
Kastovsky, D.: Word-formation, boundaries and the concept of motivation. In: Asbach-Schnitker, B. and Roggenhofer, J. (eds.), Neuere Forschungen zur Wortbildung und Historiographie der Lingustik. Festgabe für Herbert E. Brekle zum 50. Geburstag. Tübingen, 399-405.
Lieber, R.: On the Organization of the Lexicon. IULC, Bloomington 1981.
Lieber, R.: Argument linking and compounds in English. In: Linguistic Inquiry 14 (1983), 251—285.
Marchand, H.: Phonology, morphology, and word-formation. Neuphilologische Mitteilungen 1951, 87—95.
Marchand, H.: The Categories and Types of Present-Day English Word-Formation. München 1969².
Marchand, H.: Studies in Syntax and Word-Formation: Selected Articles. On the occasion of his 65th birthday, Kastovsky. D. (ed.) München 1974.
May, R.: The Grammar of Quantification. Ph. D. dissertation, M. I. T., Cambridge, Mass. 1977.
Nida, E. A.: Morphology. The Descriptive Analysis of Words. Michigan 1949².
Pesetsky, D.: Morphology and logical form. In: Linguistic Inquiry 16 (1985), 193—246.
Selkirk, E. O.: The Syntax of Words. Cambridge, Mass. 1982.
Siegel, D.: Topics in English Morphology. Ph. D. dissertation, M. I. T., Cambridge, Mass. 1974.
Strauss, S. L.: Lexicalist Phonology of English and German. Dordrecht 1982.
Trommelen, M. and Zonneveld, W.: Dutch morphology: Evidence for the right-hand head rule. In: Linguistic Inquiry 17 (1986), 147—169.
Williams, E.: On the notions "lexically related" and "head of a word" In: Linguistic Inquiry 12 (1981), 245—274.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 109–120 (1986)

ON THE ROLE OF THE AGREEMENT MORPHEME IN HUNGARIAN

I. KENESEI

(University of Szeged, Hungary)

1. Introduction

In this article I will examine a construction in Hungarian whose classification as an adjectival compound has so far been undisputed. Arguing from its distribution, the occurrence of the agreement morpheme, and non-lexically constructed phrases, I will claim that it is neither adjectival, nor a compound. Various conditions of nominative case assignment will be discussed and a comparison with similar constructions in other languages will be outlined.

2. The data and their distribution

Certain constructions usually discussed as 'adjectival subject compounds' in the Hungarian literature can apparently have two different froms with identical meanings; one (a) consists of a noun and the past participle of a verb, while the other (b) contains a noun and also a past participle, but it has a person-marking suffix as well.

(1) (a) *sors - üldöz-ött*
fate pursue PP
'pursued by fate'

(b) *sors - üldöz - t - e*
fate pursue PP Px
'idem'

(2) (a) *por - lep - ett*
dust cover PP
'dust-covered'

(b) *por - lep - t - e*
PP Px
'idem'

Other compounds of similar (a) structures cannot be converted into the person-marked (b) constructions:

(3) (a) *harc - edz - ett*
battle harden PP
'hardened by battles'

(b) ?**harc - edz - ett - e*
Px

* I wish to thank the participants at the Veszprém Conference on Morphology, May 1987, and Igor Melčuk for their comments and criticism.

(4) (a) *agy - a — lágyul - t* (b) **agy - a - lágyul - t - a*
brain-3sg soften PP Px
'softhead' (lit.: 'his-brain-softened')

(5) (a) *esz - e — vesz - ett* (b) **esz - e - vesz - t - e*
mind-3sg go PP PP Px
'desperate' (lit.: 'his-mind-gone')

Clearly, (4) and (5) differ from (1)—(3) in that they contain intransitive verbs, thus, as is usual in the literature on this question, we may suppose that the (b) type compounds can only be formed from transitive verbs. The case of (3) is somewhat more catchy: *edz* 'train, harden' is a transitive verb, but, counter to the sense-translation, it is understood as said of someone not 'hardened *by*' but 'hardened *in*' battles. Thus, strictly speaking, type (a) compounds do not all go back to subject — verb constructions.

Whereas the literature traditionally classifies all of the compounds above as adjectival—attributive, it is a highly questionable position. Certainly, adjectives can occur prenominally within noun phrases, where both type (a) and type (b) constructions are grammatical.

(6) (a)

Látt-am egy { *sors-üldözött* / *por-lepett* / *harc-edzett* / *esze-veszett* } *ember-t.*

saw-1sg an { pursued by fate / covered by dust / hardened in battles / desperate } man-ACC

'I saw a man (who was) . . .'

(b) *Láttam egy* { *sors-üldözt-e* / *por-lept-e* } *embert.*

'I saw a man { Pursued by fate.' / Covered by dust.' }

But if this prenominal position is occupied by an AP (adjective phrase) node, we may expect all of these compounds to occur also in other AP positions. That, however, is not borne out by the data.

(7) (a) *Pál (hihetetlenül)* { *sors-üldözött* / *por-lepett* / *harc-edzett* / *esze-veszett* } *volt.*

Paul-NOM inconceivably . . . . . was

'Paul was inconceivably { pursued by fate / dust-covered / hardened in battles / desperate }

(b) **Pál (hihetetlenül)* { *sors-üldözt-e* / *por-lept-e* } *volt.*

(8) (a) *Pál* { *sors-üldözött-* / *por-lepett-* / *harc-edzett-* / *esze-veszett-* } *-nek tartotta János-t.*

Paul-NOM . . . DAT regarded John-ACC

'Paul regarded John as { pursued by fate / dust-covered / hardened in battles / desperate }

(b) **Pál* { *sors-üldözt-é-* / *por-lept-é-* } *-nek tartotta Jánost.*

Note that in the positions enclosed by the curly brackets in (7) and (8), any AP can occur; it is therefore reasonable to assume that type (b) constructions are not adjectives.

3. Person-marking paradigms

Before we try to answer the question of how to categorize compounds of the (b) class, let us make a short digression into the nature of the personal endings on these constructions. Observe first of all that in the works dealing with these structures it has always been recognized that it is (or at least was, at some earlier time of its history) possible to mark them for all persons in both singular and plural, cf.:

(9) (a) *a* *tegnap* *említ-ett-em* *példa*
the yesterday mention-PP-1sg example
'the example that I mentioned yesterday'

(b)	*a*	*tegnap*	*említ-ett*-ed	*példa*	2sg
(c)			*említ-ett*-e		3sg
(d)			*említ-ett*-ük		1pl
(e)			*említ-ett*-étek		2pl
(f)			*említ-ett*-ék		3pl

It is worth noting that we have ample literature on this subject since traditional grammarians can be said to have been preoccupied with this very issue because of their belief of its being distinctive of the part-of-speech (i.e. categorial) classification of the compound. It was always the choice between the nominal and the verbal nature of the suffixation that lay in the heart of the matter. If the suffixes were more like those of the possessive paradigm, the construction would prove to be nominal, but if they were to resemble the paradigm of finite verbs, the compounds could be shown to be verbal. Here I will not dwell on the, remarkably long, history of the problem, but will simply list the forms judged as acceptable.

Hungarian marks the possessor either by the dative or the nominative, but the head noun is also marked by an agreement suffix, whose relevant morphology is as follows (harmonizing vowels are given as V wherever their actual value is immaterial):[1]

(10) *The Possessive Paradigm*

(a) Singular	(b) Plural
1. -m	-nk
2. -d	-tVk
3. -(j)$^{a}_{e}$	(j)$^{a}_{e}$/(j)$^{u}_{ü}$k

e.g. *kapu-m* 'my gate'
gate-1sg

a *fiú-k* *kapu-ja* 'the boys' gate'
the boy-pl-NOM gate-3sg/pl

az *ő* *kapu-juk* 'their gate'
the he-NOM gate-3pl
(NB. singular)

[1] For more details, see Kálmán (1985) and Kornai (1986). I will not discuss the peculiarities of this paradigm, such as the behavior of the construction with a 3pl possessor exemplified below. For a discussion of the possessive construction in Hungarian, see Anna Szabolcsi's papers in the References.

(11) *Past Tense Paradigms*[2]

(a) Definite Singular	(b) Definite Plural	(c) Indefinite Plural
1. -m	-Vk	-Vnk
2. -d	-V̄tVk	-tVk
3. -a/e	-V̄k	-ak/ek

(12) *Type (b) Compound Paradigms*

(a) Singular	(b) Plural (current)	(c) Plural (of 1939)[3]
1. -m	-Vk	-Vnk
2. -d	-V̄tVk	-tVk
3. -a/e	-a/e	-a/e
		-uk/ük

Obviously, the singular paradigms (10a), (11a), and (12a) present no problems: they are identical. The plural forms, however, show an amazing variation. (We should not be misled by (12c), which is in complete equivalence with (10b), since its bottom line is undoubtedly an artificial construct, while the first two lines are markedly archaic and/or dialectal.[4]) But this diversity should not perplex us; the individual morphological forms of person marking may legitimately vary with respect to environment. When attached to a noun, it may take a shape different from one affixed to a finite or non-finite verb. This is a phenomenon frequently encountered in the languages of the world; an example in Hungarian could be the choice of the *-i-* vs. *-k-* form of the plural affix, which depends on whether the noun is in a possessive construction or not.

Person marking can then be considered to be a unitary category whose actual inflectional form is a function of the stem (containing any possible derivational affix, e.g., infinitival, participial, etc.). Having clarified this, however, takes us no closer to answering the question of what type (b) constructions actually are.

4. The problem of compoundhood

Since compounds are a result of a lexical process, they do not contain proper names (in their referential use) or categories above zero-level in the terminology of X-bar theory. That is, whereas compounds like *student therapy* or *teacher-constructed (example)* are perfectly possible, *Kingston therapy*, on one

[2] As is well-known, verbs are marked for definite objects as against indefinite ones or the absence of an object.

[3] This is from Szepesi (1939), who cites them without giving any evidence, whether historical or other. Note that Simonyi (1907) does list data for the first three lines of (12c), but he has no examples for the notorious bottom line *-uk/ük*.

[4] Presumably Szepesi (1939) was somewhat biased to consider type (b) constructions as having possessive endings, following Simonyi (1907). He did not, however, recognize that the paradigm will thereby be identical to that of infinitives, a point that will come up in the discussion below.

hand, can only be understood as containing a non-referential mention of a certain Kingston and as referring to some *kind* of therapy rather than to a therapy carried out on or by someone called Kingston. On the other hand **a friendly teacher constructed (example)*, in which we find a non-zero level projection (*friendly teacher*), is totally ungrammatical.

These conditions are, in general, observed by Hungarian compounds too. But type (b) constructions can contain both referential proper names and non-zero level or even maximal projections.

(13) (a) *a Mari említ-ett-e példa*
the Mary-NOM mention-PP-3sg example
'the example mentioned by Mary'

(b) *a Dánia el - foglal-t-a területek*
the Denmark-NOM perf-occupy-PP-3sg territories
'the territories occupied by Denmark'

(c) *a kedves feleség-ed mutat-t-a képek*
the kind wife-your-NOM show-PP-3sg pictures
'the pictures shown by your kind wife'

(14) (a) **a Mari említett példa*

(b) **a Dánia (el)foglalt területek*

(c) **a kedves feleséged mutatott képek*

It would follow then that type (a) constructions as in (14) are indeed compounds, exactly because they are ungrammatical, while type (b) expressions are not.

Recall also that, as was illustrated in (3), certain type (a) compounds are not convertible into type (b) constructions, which was said to be due to their not containing a logical subject. Below we repeat (3a) along with an additional example.

(15) (a) *harc - edz - ett*
battle harden-PP
'hardened in battles'

(b) **harc-edz-ett-e*
3sg

(16) (a) *szél - véd - ett*
wind protect-PP
'protected from the wind'

(b) **szél-véd-t-e*
3sg

Whereas the noun in (15a) can, in principle, receive subject interpretation, that is hardly available for (16a).[5]

[5] Note that type (a) constructions are more and more shifting out of productive use. I have only been able to coin new ones of the "weather-agent" type, e.g., *vihar-tépett* 'storm-torn', *eső-mosott* 'rain-washed', *villámlás-pusztított* 'lightning-destroyed'. Still, the distinction is real: proper names and non-zero level projections are fully acceptable in type (b) constructions, but they have never surfaced in the historical data of type (a) compounds.

But the problem of whether or not it is (logical) subjecthood that is dinstinctive in (15a, b) is rendered irrelevant by the following set of examples, which show that type (b) constructions are unacceptable unless they contain an argument that has an external theta-role, i.e. an agent.

(17) (a) *a Mari szeret-t-e emberek
the Mary-NOM love-PP-3sg people
'the people loved by Mary'
(b) *a Mari tud - t - a vers
know-PP-3sg poem
'the poem known by Mary'
(c) *a könyv tartalmaz-t-a szavak
book-NOM contain-PP-3sg words
'the words contained by the book'

This requirement will also subsume the case of (3a, b), where the noun cannot be an agent, whether or not it is the subject.

We have thus arrived at the interim conclusion that type (b) constructions are not compounds (derived through some lexical rule) but syntactic categories of the subject — predicate type. If that is the case, the subject NPs within them must be marked for some case, say, nominative, according to the Case Filter, which demands that every NP with a phonetic matrix have Case.

5. Case Assignment

Nominative case has been shown to occur in two constructions in Hungarian: (a) in tensed sentences, and (b) in possessive constructions. Since the category of Tense does not play any role in the latter, Case Assignment must be dependent on the AGR constituent of INFL.

(18) (a) Te tud-t-ad a vers-et.
you(sg)-NOM know-Past-Def. 2sg the poem-ACC
'You knew the poem.'
(b) a te ház-ad
the you-NOM house-sg
'your house'

Whereas in tensed sentences the subject need not be placed next to the inflected verb, as this position is reserved for the focus, some adjacency requirement is in force in possessive constructions. If the head noun, which is marked for agreement, is separated from the specifier by the definite article, the nomi-

native case is unavailable for the possessor, and it must be marked dative. Note that if the possessor is outside the NP containing the head, which is a standard discontinuous construction in Hungarian, it must also be in the dative.

(19) (a) *Olvastam* [*János* *könyv-é-t*]
I-have-read John-NOM book-3sg-ACC
'I have read John's book.'
(b) **Olvastam* [*János a könyvét*]
the
(c) *Olvastam* [*János-nak a könyvét*]
DAT
(d) *Jánosnak olvastam a könyvét*

It is also worth mentioning that a third context in which person marking occurs allows subjects only in the dative. The infinitial constructions given below can in general be split up and their constituents placed anywhere in the matrix clause. Arguably, there are no infinitival clauses at S-structure in Hungarian, but the discussion of this issue would take us too far afield.

(20) (a) *János-nak* *Péter-rel* *kell-ett* *beszél-ni-e.*
John-DAT Peter-with must-Past speak-Inf-3sg
'John had to speak with Peter.'
(b) *Nek-tek* *kár* *volt* *dolgoz-no-tok.*
DAT-you(pl) no-use was work-Inf.-2pl
'It was no use for you to work.'

Thus it seems that there are two prerequisites for a noun phrase to receive nominative case: (1) the head (i.e. V or INFL in S, N in NP) must be marked for agreement, and (2) the NP must be adjacent to the head (except in tensed clauses). The constructions discussed here observe both criteria: the non-finite verbs are affixed for agreement and the subject must not be separated from the verb by any material not part of the lexical entry for the verb.

(21) (a) *a* *Mari* (*meg-*) *vizsgál-t-a* *betegek*
the Mary-NOM perf. examine-PP-3sg patients
'the patients examined by Mary'
(b) **a Mari tegnap* (*meg*)*vizsgálta betegek*
yesterday

[6] We are not concerned here with accounting for the often intriguing behavior of possessive constructions in Hungarian. For some discussion, see in addition to Szabolcsi's work, Kornai (1989) and Kenesei (1986).

The ungrammaticality of (21b) is comparable to that of a tensed sentence in which the (would-be) focus is separated from the inflected verb by some other constituent.

(22) (a) *Tegnap Mari vizsgálta meg a betegeket.*
yesterday Mary-NOM examined Perf. the patients-ACC
'It was Mary that examined the patients yesterday.'
(b) **Mari tegnap vizsgálta meg a betegeket.*

(Note that (22b) is starred only if *Mari*, rather than *tegnap*, is in focus.)

A few further problems still remain. For example, we have no answer to the question of why there cannot be type (b) constructions of more than two constituents: usually subject plus verb, cf.:

(23) **a tegnap Mari (meg-)vizsgálta betegek*
the yesterday Mary-NOM Perf.-examined patients
'the patients examined by Mary yesterday'

And it need not always be the subject that fills in for the preverbal constituent; in other than third person forms, as is usual in this pro-drop language, personal pronouns can be omitted. But then the occurrence of another preverbal constituent seems to be obligatory.

(24) (a) *a* ?*(*most*) *említ-ett-em* *példa*
the now mention-PP-1sg example
'the example I just mentioned'
(b) *az* ?*(*imént*) *idéz-t-ük* *mondat*
the just quote-PP-1pl sentence
'the sentence we just quoted'

Certainly, person marking in other than third persons is rather archaic or awkward and has been replaced by the use of the paradigm of the emphatic/reflexive pronouns *magam* 'myself', *magad* 'yourself', etc., which invariably trigger agreement in the third person both here and in the possessive constructions (contrary to tensed sentences, where it requires ordinary agreement)

(25) (a) ?*a tegnap készít-ett-em kép*
the yesterday make-PP-1sg picture
'the picture I made yesterday'
(b) *a magam készít-ett-e kép*
myself-NOM 3sg
'the picture I made (myself)'

(c) *a magam kép-e/*kép-em*
3sg 1sg
'my (own) picture/a picture of myself'
(d) *A képet magam készít-ett-em/*-e.*
the picture-ACC myself-NOM make-Past-1sg/3sg
'I made the picture myself.'

6. Categorization

As was seen in the foregoing sections, the constructions discussed are best treated as non-finite clauses. We may suppose that in Hungarian, as in a number of other languages (mainly of the SOV type), there is a prenominal S position within the NP to be filled in either by finite clauses (as in Japanese, Korean, Sinhalese) or by non-finite ones (as in most Altaic and a large number of Uralic languages). That is why these clauses cannot occur in the position of ordinary adjectives; on the one hand, they are not adjectival, and, on the other, they contain an empty category that is left uncontrolled if the clause is outside the NP.

Observe that other, allegedly adjectival, constructions also exhibit the behavior illustrated in (7) and (8), i.e., they cannot be classified as adjectives on distributional grounds.

(26) (a) *a betegeket vizsgáló orvos*
the patients-ACC examining doctor
'the doctor examining the patients'
(b) **Mari betegeket vizsgáló volt/maradt*
Mary was remained
(27) (a) *a tegnap megvizsgált betegek*
the yesterday examined patients
'the patients examined yesterday'
(b) **A beteg [tegnap megvizsgált] volt*

Note also that since the non-finite verb has no person-marking (i.e. AGR), no nominative subject/agent can occur in (27a), although an agent marked by a specific agentive case-suffix or postposition is possible.

(28) (a) **a Mari tegnap megvizsgált betegek*
the Mary-NOM yesterday examined patients
(b) *a* Mari { -tól / által } tegnap *megvizsgált betegek*
by
'the patients examined by Mary yesterday'

7. Nominative case in non-finite clauses in other languages

Some of the languages that have prenominal non-finite clauses apply person-marking on the verb, but we have no reason to suppose that the subject is in a case other than genitive, cf.:

(29) Turkish (source: Sezer (1986))
(a) *sen-in gör-düğ-ün filim*
you-GEN see-Part-2sg film
'the movie that you saw'
(b) *al-dığ-ı araba*
buy-Part-3sg car
the car he bought

(30) Finnish (source: T. Mikola (personal communication))
(a) *isä-n teke-mä tuoli*
father-GEN make-Af chair
'the chair the father made'
(b) *Isä istun teke-mä-llä-nsä tuoli-lla.*
father-NOM sat make-Afx-Case-3sg chair-Superess.
'The father sat on the chair he had made.'

Note that while Turkish has obligatory person marking all along, Finnish makes use of it only in case there is an (overt or covert) personal pronoun in the subject/possessor position.

Ostyak, a Finno-Ugric language, and Evenki, another Siberian language of the Tungusic family, resemble the pattern seen in Hungarian more closely than Turkish or Finnish.

(31) Ostyak (source: Hajdú (1973))
mä wəmam weli kalas
I-NOM bought-1sg reindeer perished
'The reindeer I bought has perished.'

(32) Evenki (source: Comrie (1981))
bi pis'mo-wa ga-ća bi-si-m akii-m
I letter-ACC receive-PastPart. be-Pres-1sg brother-1sg
min-dulə unʒə-rii-wəə-n
I-LOC send-Part-ACC-3sg
'I have received the letter which my brother sent to me.'

In (31), the head noun is preceded by the person marked participle and the personal pronoun in the nominative. In (32), the sentence final participle is related to its head by the accusative affix on the verb form, which agrees in person and number with the nominative subject of the clause.

8. Conclusion

It has been shown that what was called;"type (b) constructions" are non-finite clauses within NPs. That is why they allow person marking (i.e. AGR) to occur, which in turn makes it possible to assign nominative case to the subject under certain conditions. They conform to the behavior of non-finite clauses in this language and resemble corresponding structures in other languages.

References

Comrie, B.: The Languages of the Soviet Union. Cambridge University Press, Cambridge 1981.

Hajdú, P.: Bevezetés az urális nyelvtudományba [Introduction to Uralic linguistics]. Tankönyvkiadó, Budapest 1973.

Kálmán, L.: Inflectional Morphology. In: Kenesei, I. (ed.): Approaches to Hungarian 1. JATE, Szeged 1985, 247-262.

Károly, S.: Az *istenadta*-féle szerkezetek személyragos tagjának szófaji jellegéről [On the part-of-speech classification of the person-marked constituent of type (b) consctructions]. In: Nyelvtudományi Közlemények 59 (1957), 130—150.

Kenesei, I.: Az alárendelő mondatok [Subordinate clauses]. In: Nyelvtudományi Közlemények 88 (1986), 311—350.

Kornai, A.: On Hungarian Morphology. Unpublished doctoral dissertation, Institute for Linguistics, Budapest 1986.

Kornai, A.: A főnévi csoport egyeztetése [Agreement in the noun phrase]. In: Általános Nyelvészeti Tanulmányok XVII, 1989, 183-211.

Sezer, E.: Configurationality, Proper Government alla Turca, and Don't Mova Alpha, paper read at the Workshop on AUX, Configurationality and Lexical Structure. Groningen—Amsterdam 1986.

Simonyi, Zs.: Igenévi szerkezetek [Non-finite constructions]. Nyelvészeti Füzetek 47, Athenaeum, Budapest 1907.

Szepesi, Gy.: Az *istenadta*-féle szerkezetek a finnugor nyelvekben [Type (b) constructions in the Finno-Ugric languages]. Magyar Nyelvtudományi Társaság, Budapest, 1939.

Szabolcsi A.: The possessive construction in Hungarian. In: ALH 31 (1981), 261—289.

Szabolcsi A.: The possessor that ran away from home. In: The Linguistic Review 3 (1983), 89—102.

Szabolcsi A.: A birtokos szerkezet és az egzisztenciális mondatok [The possessive construction and existential sentences]. Unpublished doctoral dissertation, Institute for Linguistics, Budapest 1986.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 121–133 (1986

LEXICAL MECHANISMS VERSUS MORPHOLOGICAL STRUCTURE

J. VAN MARLE*

(P. J. Meertens-Instituut, The Netherlands)

1. Introduction

Of late quite a number of linguists have tried to work out the general idea that languages should be regarded as systems in which a large number of — interacting and opposing forces — contend. Languages, in this view, are by definition in an unsteady equilibrium: they always involve some sort of compromise between the diverse forces they comprise. Crucial to this — interactionist — conception of language is also, that the pressure to achieve linguistic optimality is nearly always considered "local", i.e. it can be seen to concentrate on one part of the system while leading to less desired results in other parts. Put differently, due to the multi-faceted character of language, the pressure to linguistic optimality nearly always focuses on one aspect/component of the language, which implies that other kinds of linguistic optimality (bearing upon other facets of the language) often deteriorate.[1]

With respect to word-formation the above conception of language has been elaborated in so-called n a t u r a l m o r p h o l o g y, a trend which, since its very beginning, involves such an interactionist model (Dressler, 1986). In their search for forces which are in conflict with the principles of word-formation — and which, consequently, may affect optimal morphological structure —, the adherents of natural morphology have rightly stressed that particularly phonetics/phonology is a domain of language which accomodates many forces undermining the transparency of morphology. No doubt Mayerthaler is correct when he characterizes phonology as k o n t r a m o r p h o l o g i s c h (i.e. counter-morphological), which means that in his view phonology and morphology (apart from just a few exceptions) can never be optimalized at the same time (Mayerthaler, 1981, 43).

* I am indebted to the participants of the conference for their comments and criticisms, in particular to G. E. Booij, R. P. Botha, and W. U. Dressler. Moreover, I am indebted to W. Meijs who checked my English

[1] General studies in which this conception of language is elaborated are Koefoed (1974; 1978) and Dik (1986). Cf. also Dressler (1986) (with many further references) and Mayerthaler (1981) for morphology, Langacker (1977) for syntax, and Van Marle (1978) for an attempt to interpret morphological change in terms of two opposing forces

Particularly in the work of Dressler attention is drawn to the fact that phonetics/phonology is not the only domain of language which may be in direct conflict with word-formation. Another component of language, as is rightly stressed in Dressler (1977), which may interfere with the principles of morphology is t h e l e x i c o n. The central topic of this paper, then, is the discrepancies/conflicts between lexicon and morphology, as I hold the view that the discrepancies between these two levels have received too little attention in the past few years.[2]

From the above it follows that in my opinion a distinction should be made between morphology and the lexicon. Morphology deals with the systematics of c l a s s e s o f w o r d s — or, as I prefer to think of it, of w o r d - t y p e s —, whereas the lexicon deals with properties of individual words (Van Marle, 1985, ch. 3). As said, the introduction of this distinction means that I start from the idea that there may be discrepancies between the word viewed as a direct representative of a given word-type and the word as a separate, individual and 'unique' entity, i.e. as a sign. To a large extent, such discrepancies between the morphological and the lexical level result from the fact that in the words stored in the lexicon the efficacy of l e x i c a l m e c h a n i s m s may be present. These lexical mechanisms may be in conflict with the principles of word-structure, which means that they may affect optimal morphological structure and/or render it opaque.

Logically three types of discrepancies between the morphological and the lexical level may be distinguished and, interestingly, all three of them actually occur. These discrepancies are:

(1) (a) morphological difference corresponds to lexical identity;
(b) morphological identity corresponds to lexical difference; and
(c) morphological difference corresponds to lexical difference, but the nature of both differences is not the same.

In the remainder of this paper the above three discrepancies between the morphological and the lexical level will be elaborated. The first two cases will be dealt with only cursorily, whereas the third case will be discussed in some more detail.

2. Absence of lexical opposition

The first case to be discussed is the situation in which a distinction on the morphological level does not correspond to a distinction on the lexical level. Specifically, the difference in meaning between two morphological

[2] Note, however, that I will n o t be concerned with the well-known phenomenon of lexicalization

categories is not always realized on the lexical level, i.e. the difference in meaning is not always reflected by the individual words representing the word-types in question. That is:

(2) "Difference" on the morphological level corresponds to "identity" on the lexical level.

On a general level of theoretical reflection discrepancies of this kind between the morphological and the lexical level are well-known, of course. Typically, discrepancies of this kind are generally considered manifestations of the phenomenon of morphological neutralization. Note, however, that, as far as word-formation is concerned, many aspects of morphological neutralization are unclear, while it is hard to avoid the impression that the phenomena classed under the heading of morphological neutralization constitute a rather heterogeneous set.[3] All in all, the phenomenon of morphological neutralization deserves a separate treatment, which means that within the context of the present paper I can only make some brief, and necessarily tentative, remarks with respect to one specific type of neutralization: the neutralization taking place in bipartite schemes of modification.

It is my impression that, as far as word-formation is concerned, morphological neutralization is particularly prominent in the domain that is usually referred to as modification. Moreover, the neutralization of the semantic differences between two modifying word-types seems to be subjected to the general proviso that the two categories must, as it were, express two different gradations of what from a semantic point of view seems to be essentially one and the same process.

Put in general terms, such bipartite systems of modification result in triplets (including the base) of the following format:

(3)

base		modification 1	modification 2
(B)	↖	(M 1)	(M 2)

In his early structural sketch of the Russian noun, Karcevski encountered such a bipartite system of modification in the domain of diminuation. Beside the base, there is both the "simple diminutive" and the "lauditive-diminutive" (the terminology is inexact), cf. *les-lesók-lesóček* (Karcevski, 1932, 342).

[3] E.g. a distinction should be drawn between the neutralization of the semantic differences between a derivative and its base, and the neutralization of the semantic differences of two related word-types (it is only the latter type of neutralization that is discussed below). In addition, in certain cases neutralization is clearly restricted to a specific — syntagmatic — context, whereas other types of neutralization are not governed by syntagmatic forces

At least as early as Karcevski (1932) it is also known, however, that triplets of the format of (3) are not infrequently upset, which means that there may be question of "déplacements" (i.e. transfers) among their members. As to the case of diminuation in Russian, Karcevski notes (i) that in some cases the "simple diminutive" has become "un mot ordinaire", i.e. has lost its diminutive character, whereas (ii) the — original — base has become some sort of "augmentative" (ibid., 342). Crucial, then, to the topic of this paper is, that in a l l cases in which a representative of M 1 has lost its diminutive character, a representative of M 2 takes its place immediately and turns into a "simple diminutive". That is, if a representative of M 1 drops out of the bipartite system of modification and its counterpart of M 2 takes its place, the bipartite system of the morphological level corresponds to a simple system of diminuation on the lexical level.

Clearly, the above transfer of M 2 to M 1 illustrates the discrepancy between morphological and lexical level that constitutes the subject-matter of this section. On the morphological level two modifying categories should be distinguished. However, on the lexical level — i.e. in the individual words —, we cannot but ascertain that the difference distinguishing both word-types is not always present. That is, some words have the formal characteristics of M 2, whereas they have the semantic properties of M 1 (viz. when the proper representative of M 1 has lost its diminutive character). On the lexical level, to put it differently, the semantic difference between both word-types disappears when the diminutive character of M 1 is affected, i.e. when the representatives of both modifying categories which share the same base are never in lexical opposition.[4]

There can be no doubt that the type of morphological neutralization sketched above renders the systematics of the morphological system opaque. The reason is that the notion 'simple diminutive' is sometimes expressed by a representative of M 1 — i.e. when the diminutive character of this formation is not affected —, and sometimes by a representative of M 2 — i.e. when the

[4] It may well be the case that there are differences among bipartite systems of modification in the way(s) in which such systems may be affected. However, the general trend discussed above — i.e. the strong tendency for the representatives of the marked category to occupy the 'open' position of the representatives of their unmarked counterpart — seems to be well-established. Put differently, l e x i c a l o p p o s i t i o n i s a n e c e s s a r y c o n d i t i o n t o t h e f u n c t i o n i n g o f s u c h b i p a r t i t e s y s t e m s o f m o d i f i c a t i o n.

Cf. Schultink (1962, ch. 3) for a detailed discussion of the bipartite system of deadjectival relativization in Modern Dutch. Besides simplex *groen* 'green' we find in this language both *groen-ig* and *groen-erig*, the difference between both forms being a matter of degree: formations in *-erig* display a higher degree of relativization than their counterparts in *-ig* (which, of course, completely concurs with general semiotic principles). However, in those cases where the formations in *-ig* are systematically absent (i.e. where there is question of a systematic gap), or where they lack a relativizing character, the formations in *-erig* take their place immediately (which implies that in these cases

diminutive character of its counterpart of M 1 is affected. The transfer of the representative of M 2 to the position of its counterpart of M 1 (in combination, of course, with the lexicalization of the representative of M 1) seriously affects the trend of b i u n i q u e n e s s which is one of the main features of the systematics of the word-formation component.

3. 'Accidental' lexical opposition

The second case I wish to discuss is the situation in which similarity on the morphological level does n o t correspond to similarity on the lexical level. Or:

(4) "Identity" on the morphological level corresponds to "difference" on the lexical level.

As far as I am aware, this situation is particularly common in those cases in which a morphological system gives rise to m o r p h o l o g i c a l d o u b l e t s, i.e. when there are rival morphological categories.

In a way, morphological doublets may be looked upon as the result of the fact that we have to do with a, to a greater or lesser extent, ill-balanced system. Put differently, 'ideally' morphological systems should not give rise to doublets at all; morphological doublets arise in systems which lack a cast-iron structure.[5]

Morphological doublets come about in systems in which there are rival morphological categories which overlap. Rival morphological categories display the same semantic characteristics, which implies that on the morphological level there is no question of opposition; rival morphological processes are essentially complementary. On the lexical level, however, things are different. If rival morphological processes overlap, i.e. (may) generate doublets, the resulting doubles are in opposition on the lexical level. Crucial to doublets is, then, that the lexical opposition they display — due to the fact that they are generated by rival morphological processes — does n o t correspond to (or: is

---

the words in *-erig* do not express the high degree of relativization that their congeners which are in — potential — opposition with a counterpart in *-ig* do).

It may well be the case that the famous 'double diminutives' in Afrikaans (Odendal, 1963) bear upon this issue as well, regardless of the fact that there is no unanimity as to the details of their semantics (Botha, 1984, 174, fh. 48). However the details of their semantics may be, these 'double diminutives' display their specific semantic characteristics — which Odendal (1963, 222) defines as "exceptional smallness or insignificance or excessive sympathy, tenderness, contempt, disgust and the like" — only, if they are in — potential — opposition with fully transparent "simple diminutives". If the diminutive character of the "simple diminutives" is affected (which is a very common trend in Afrikaans), the "double diminutives" lose their specific characteristics immediately and adopt the characteristics for "simple diminutives"

[5] No doubt, this remark is in need of further refinement. Cf. Van Marle (1985, part 3; 1986) for more subtleties concerning rival morphological processes and the types of doublets that such systems may result in

n o t supported by) opposition on the morphological level. To put it differently, from a strictly morphological point of view the lexical opposition manifested by doublets is completely "accidental".

Modern Dutch plural-formation nicely illustrates a morphological system which is ill-balanced in the above sense. In this language there are two productive plural-categories, viz. the "special case" in *-s* and the "general case" in *-en* (cf. Van Marle, 1985, ch. 6—7). Details apart, the derivational domain of the special case in *-s* consists of a restricted number of either phonologically or morphologically determined sets of words, some of which are listed in (5):

(5) (a) words ending in *əl*, *əm*, *ən* and *ər*;
(b) words in *-aar*, *-erd* and *-aard*;
(c) some sets of words that are vowel-final, e.g. most words ending in *ə*.

Although there can be no doubt that there is a strong tendency for words belonging to the sets specified in (5) to have *-s* plurals, there is no question of a cast-iron system. All kinds of irregularities occur, only some of which will be dealt with below. Consider (6) and (7) in which I have listed some words belonging to the sets specified in (5) but which, nevertheless, have a plural in *-en*:

(6) (a) *apostel* 'apostle', *christen* 'Christian', *discipel* 'disciple', *engel* 'angel', *heiden* 'heathen', *schepen* 'sheriff, alderman';
(b) *hoogleraar* 'professor', *leraar* 'teacher';
(c) *gedachte* 'thought', *gelofte* 'vow', *weduwe* 'widow';

and

(7) (a) *bestuurder* 'governor, manager', *pakhuismeester* 'warehousekeeper', *tafel* 'table', *vader* 'father';
(b) *bastaard* 'bastard', *beoefenaar* 'student, practitioner', *eigenaar* 'owner';
(c) — *bende* 'gang, band', *rente* 'interest', *ziekte* 'illness';
— *lelie* 'lily'.

To a large extent, the formations in (6) and (7) are remnants of an earlier period of Dutch in which *-s* plurals were far less popular than in Modern Dutch. The difference between the above two sets of words is as follows:

(8) (a) the *-en* forms of the words in (6) are the 'normal', i.e. stylistically unmarked, plurals of the words in question; whereas
(b) the *-en* forms of the words in (7) are n o t the 'normal', i.e. stylistically unmarked, plurals of the words in question.

This difference between the words in (6) and (7) turns out to be crucial. The *-en* plurals of the words in (6) radiate a strong blocking force with respect to their "regular", 'to be expected' counterparts in *-s*, which means that in most cases these parallel plurals in *-s* do not regularly occur. When they are coined, however, they display a strong colloquial or even sub-standard nature. The *-en* plurals of the words in (7), on the other hand, are fundamentally different. These *-en* plurals are parallelled by stylistically unmarked *-s* plurals without exception, whereas the *-en* plurals themselves have a strong 'solemn, dignified, or even archaic' connotation. That is, for the words in (7) it is the *-s* plurals that are the "normal", i.e. stylistically unmarked, plural-forms.

The above facts give rise to the following hypothesis as to the precise stylistic characteristics of the doublets Modern Durch plural-formation gives rise to:

(9) If a word has a plural in *-en*, whereas — according to the living structure of Dutch — it should have a plural in *-s*, there are two possibilities:
  (a) If the *-en* plural is stylistically unmarked, the parallel plural in *-s* — if coined at all — is highly colloquial or even sub-standard (which means that most of the times these *-s* plurals are avoided);
  (b) If the *-en* plural is parallelled by a stylistically unmarked *-s* plural, the *-en* plural displays a solemn, dignified or even archaic character itself.

Crucial to the stylistic connotations that Dutch plural-forms according to (9) may display is, that they are "normally" absent. That in (i) "normally" *-s* plurals are neither colloquial nor sub-standard whereas (ii) "normally" *-en* plurals are neither solemn, dignified nor archaic. There stylistic features arise only in doublets, i.e. when there is question of l e x i c a l o p p o s i t i o n. From this it follows that the mechanisms of stylistic differentiation which we see at work here, are purely lexical: "normally" plural-forms (either in *-en* or *-s*) are stylistically unmarked, which is to say that stylistic differentiation crops up only when there is a question of opposition on the lexical level. Clearly, stylistic differentiation can best be regarded as a specific manifestation of the general l e x i c a l trend t o a v o i d s y n o n y m y.

Note, finally, that the mechanisms of stylistic differentiation affect the transparency of the morphological system and that their effects run counter to the general morphological trend of biuniqueness. The mechanisms of stylistic differentiation have as their effect that both a minority of the *-en* plurals and a minority of the *-s* plurals is stylistically marked, whereas, from a purely morphological point of view, both the set of stylistically marked *-en* plurals and the set of stylistically marked *-s* plurals display a completely accidental make-up.

4. The rise of equipollent oppositions

The third and final discrepancy between the morphological and the lexical level that I wish to discuss is as follows: a morphological distinction corresponds to a distinction on the lexical level, but the distinction on the lexical level differs from the distinction on the morphological level. That is:

(10) "Difference" on the morphological level corresponds to "difference" on the lexical level, but both differences are not identical.

Consider the following discussion which will, hopefully, clarify the above.

Modern Dutch has two intimately related, productive categories of deverbal personal names, viz. one in *-er* and one in *-ster*. The latter category bears upon 'specifically female' persons, whereas the former category is sex-neutral, i.e. relates to both women and men. Consider:

(11)	verb		neutral	specifically female
	frunnik-en	'to fumble'	*frunnik-er*	*frunnik-ster*
	snoep-en	'to eat sweets'	*snoep-er*	*snoep-ster*

As said, the formations in *-er* are sex-neutral. This is evidenced by the fact that these formations can be used in relation to both women and men, as can be seen from the following sentences:

(12) { *Charles* / *Diana* } *is een echte frunnik-er.*

'C/D is a real fumbler.'

and

(13) { *Ronald* / *Nancy* } *is een echte snoep-er.*

'R/N is a real sweet-eater.'

However, not all formations in *-er* conform to these patterns. Formations such as *handball-er* 'hand-ball player', *verpleg-er* 'nurse' or *zwemm-er* 'swimmer' display a behaviour that deviates from that of *frunnik-er* and *snoep-er* discussed above. Consider the following sentences which illustrate the difference between both sets of words:

(14) { ??? *Paul* / *Linda* } *is een echte handball-er.*

'P/L is a real hand-ball player.'

(15) { ??? *Mick* / *Bianca* } *is een echte verpleg-er*

'M/B is a real nurse.'

(16) { ??? *David* / *Angie* } *is een echte zwemm-er.*

'D/A is a real swimmer.'

Clearly, crucial to the formations in *-er* in question is, that they cannot — or only with great difficulty — be used with reference to women, whereas they can be used in relation to men without difficulty. To my mind, this can only be interpreted as an indication of the fact that *handball-er*, *verpleg-er* and *zwemm-er* have lost — or: are in the way of losing — their neutral character, i.e. they have become — or: are in the way of becoming — masculine.

What should be stressed is, that there can be no doubt as to the fact that *frunnik-er* and *snoep-er* represent the regular, "living" structure of Modern Dutch, whereas *handball-er*, *verpleg-er* and *zwemm-er* display secondary semantic developments. This judgement is based on the fact that newly-coined words in *-er* in Dutch display the pattern exemplified in (12) and (13) without exception, and it is the newly coined words which I consider to give the most accurate picture of the systematics actually in force (this 'rule of thumb' is called Schultink's principle in Van Marle (1985, 70, fn. 13)). That is, when I coin the personal name in *-er* of verbs such as *kot-en* 'to play at knuckle-bones' or *toep-en* 'to play at cards' — i.e. verbs which I have taken from the dictionaries and which were hitherto unknown to me —, the resulting *kot-er* and *toep-er* display a clear-cut sex-neutral character as is evidenced by (17) and (18):

(17) { *Dante* / *Beatrice* } *is een echte kot-er.*

and

(18) { *Romeo* / *Julia* } *is een echte toep-er.*

No doubt, the above gives rise to the question what is the cause of the transition from neutral to male that *handball-er*, *verpleg-er* and *zwemm-er* display. As to the answer to this question I start from the following hypothesis:

(19) It is only the formations in *-er* which are parallelled by a counterpart in *-ster* that is c u r r e n t[6] — i.e. which has become part of the lexical stock of the language —, which undergo the transition from neutral to male.

If, in other words, a formation in *-er* lacks a counterpart in *-ster* that is current (and most newly coined formation in *-er* do !), this formation is and remains neutral, however current it may be itself!

Given the above, the discrepancy between the morphological and the lexical level can be looked upon as follows: on the morphological level the categories (the word-types) in *-er* and *-ster* stand in a p r i v a t i v e o p p o s i t i o n : the marked category in *-ster* must be defined positively, whereas the unmarked category in *-er* is to be defined negatively, i.e. as 'n o t specifically female'. On the lexical level, however, the relationship between a formation in *-er* and its counterpart in *-ster* may be different: if the latter formation is current, the relationship between both formations is (or tends to be) that of equipollency, as both terms of the lexical opposition are to be defined positively, viz. one as 'male' and one as 'female'.

What seems to be the case, then, is that, if a formation in *-ster* is current, its — originally neutral — counterpart in *-er* is conceived of as the d i r e c t o p p o s i t e of the formation in *-ster*.[7] This reinterpretation that the formations in *-er* may be subjected to, leads to the change from a privative opposilion on the morphological level to an equipollent opposition on the lexical evel. As to the nature of the mechanisms underlying the reinterpretation of the originally neutral formations in *-er* as the direct opposites of the formations in *-ster* that are current, I would venture to formulate the hypothesis that we have to do here with an instance of the trend towards l e x i c a l o p t i m a l i t y.

Lexical optimality bears upon the "ideal" relationships between the words stored in the lexicon. Consequently, lexical optimality bears upon the relationships between words, viewed as separate, independent entities. Clearly, this grouping of words into semantic configurations — into W o r t f e l d e r (wordfields) — need by no means be in harmony with the systematics of word-formation, in consequence of the fact that, as Dressler (1977, 17) puts it, "word-fields group words irrespective of their compositional or non-compositional character". As to the phenomena discussed in Lyons (1977) under the

[6] Note that words which are "current" need not be characterized by unpredictable formal and/or semantic properties.

[7] It may well be the case that the transition from neutral to male that the formations in *-er* may undergo forms part of a more general phenomenon. Cf. Lyons' observation that certain types of antonyms "are frequently employed in everyday language-behaviour as contradictories rather than contraries. If we are asked *Is X a good chess-player?* and we reply *No*, we may well be held by the questioner to have committed ourselves implicitly to the proposition that *X* is a bad chess-player".

heading of 'opposition and contrast', I start from the hypothesis that, on the lexical level, the optimal relationship between the words in question is such that they constitute bipartite sets for which it holds that both words are each other's d i r e c t o p p o s i t e. Both words are of the same value, which means that the relationship between the members of such sets is that of equipollency.[8]

Evidence in favour of this view is, that — as far as the domain of 'opposition and contrast' is concerned — among simplex words equipollency is highly current. To stick to the field of sex-differentiation, cf. the words listed in (20) which constitute such sets of male and female counterparts without exception:

(20)	I		II	
(a)	*kip*	'chicken'	*haan*	'cock'
	koe	'cow'	*stier*	'bull'
	merrie	'mare'	*hengst*	'stallion'
	ooi	'ewe'	*ram*	'ram'
	poes	'pussy-cat'	*kater*	'tom cat'
	teef	'female dog'	*reu*	'male dog'
(b)	*dochter*	'daughter'	*zoon*	'son'
	meisje	'girl'	*jongen*	'boy'
	moeder	'mother'	*vader*	'father'
	nicht	'niece'	*neef*	'nephew'
	tante	'aunt'	*oom*	'uncle'
	vrouw	'woman'	*man*	'man'

In consequence of their transition from neutral to male, the relationship between *handball-er*, *verpleg-er* and *zwemm-er* and their specifically female counterparts *handbal-ster*, *verpleeg-ster* and *zwem-ster* strongly resembles the relationship between the words in (20) II and their counterparts in (20) I, viz. the relationship of a male person/animal to its direct — female — opposite.

Another indication in favour of my claim that it is an instance of the trend towards lexical optimality that we have to do with here is, that the transition from neutral to male described above is by no means restricted to the formations in *-er*. It is striking to see, for instance, that the relationship between *typist* 'id.', *agent* 'policeman', *echtgenoot* 'husband' and *leerling* 'pupil' and their specifically female counterparts in *-e* (i.e. *typist-e*, *agent-e*, *echtgenot-e* and *leerling-e*) seems to be identical to the relationship between the formations

[8] Clearly, this boils down to the claim that — at least, in certain parts of the lexicon — equipollent oppositions are most natural. Many aspects of the optimal lexical structure require further investigation. It may well be the case that Trubetzkoy's general observation that in any system equipollent oppositions are most frequent (Trubetzkoy, 1939, 67), is of interest in this connection

in *-ster* that are current and their counterparts in *-er*, regardless of the crucial morphological difference between the two sets of words. (The female formations in *-e* have their originally neutral counterparts as their derivational base, whereas the formations in *-ster* are *not* derived from their counterparts in *-er*.) Words in *-ist*, *-ent* or *-ant*, *-genoot*, or *-ling* which lack a female counterpart in *-e* that is current, are neutral; however, in those cases in which these formations do have a female counterpart in *-e* that is current (such as the examples presented above), the transition from neutral to male can be observed in these words as well.[9]

If my above suggestion that the change from neutral to male can also be found in other — from a morphological point of view, crucially different — cases is correct, it directly supports my claim that it is an instance of the trend towards lexical optimality that we tracked down in the above, since it turns out to be operative irrespective of the properties of morphological structure.

With the above discussion of the transition from neutral to male I hope to have illustrated the third discrepancy between the morphological and the lexical level. A privative opposition on the morphological level corresponds to an equipollent opposition on the lexical level. Due to the efficacy of the trend towards lexical optimality, the originally neutral formations in *-er* are conceived of as the direct opposites of the female counterparts in *-ster* that are current. That is, the formations in *-er* are subjected to a lexical arrangement which, crucially, is fundamentally different from morphology. At least in certain domains of the lexicon equipollency seems to prevail such that it may accomplish that the morphological trend towards biuniqueness gets blurred: some formations in *-er* are neutral, whereas other formations in *-er* are male. Note, finally, that in this case, too, the set of formations in which the effects of lexical mechanisms is present displays a make-up which from a morphological point of view is completely accidental.

5. Concluding remarks

In the preceding I discussed three types of lexical facts which are at the root of the three discrepancies between the lexical and the morphological level distinguished in the opening section of this paper:

(21) (a) the absence of lexical opposition;
(b) lexical opposition which does not correspond to morphological opposition; and
(c) a preferred lexical opposition-type.

[9] Whether all Dutch personal names join in with the patterns specified above — and if so, to the same extent — is a matter for further investigation

Two remarks seem to be in order. In the first place it should be remarked that insight into the effects of lexical mechanisms deepens our insight in morphology. That is, I consider the study of lexical mechanisms instrumental in unearthing the systematics of morphology proper, as it is only by taking the effects of the 'obliterating' lexical forces into account that the precise structuring of morphological systems — and the properties of the words they underlie — can be fully understood. Secondly, it seems noteworthy to emphasize that the lexical mechanisms discussed above by no means display the unsystematic and haphazard nature that lexical phenomena are so often associated with.

References

Botha, R. P.: A Galilean Analysis of Afrikaans Reduplication. Stellenbosch Papers in Linguistics 13 (1984).
Dik, S. C.: On the Notion 'Functional Explanation'. Working Papers in Functional Grammar (1986).
Dressler, W. U.: Elements of a polycentristic theory of word formation. Wiener Linguistische Gazette 15 (1977), 13—32.
Dressler, W. U.: Explanation in natural morphology. Linguistics 24 (1986), 519—548.
Karcevski, S.: Sur la structure du substantif russe. Reprinted in: Vachek, J. (ed.): A Prague School Reader in Linguistics. Indiana University Press, Bloomington 1964 335—346. [1932]
Koefoed, G. A. T.: On formal and functional explanation: some notes on Kiparsky's "Explanation in phonology". In: Anderson, J. M. and Jones, Ch. (eds.): Historical Linguistics II. North-Holland, Amsterdam 277—293.
Koefoed, G. A. T.: Taalverandering in het licht van taalverwerving en taalgebruik. In: Koefoed and Marle (eds.), 1978, 11—70.
Koefoed, G. A. T. and Marle, J. van (eds.): Aspecten van taalverandering. Wolters-Noordhoff, Groningen 1978.
Langacker, R. W., Syntactic reanalysis. In: Li, Ch. N. (ed.): Mechanisms of Syntactic Change. University of Texas Press, Austin 57—139.
Lyons, J.: Semantics I. Cambridge University Press, Cambridge 1977.
Marle, J. van: Veranderingen in woordstructuur. In: Koefoed and Marle (eds.), 1978, 127—176.
Marle, J. van: On the Paradigmatic Dimension of Morphological Creativity. Foris, Dordrecht 1985.
Marle, J. van: The domain hypothesis: the study of rival morphological processes. Linguistics 24 (1986), 601—628.
Mayerthaler, W.: Morphologische Natürlichkeit. Athenaion, Wiesbaden 1981.
Odendal, F. F.: Limitations of morphological processes. A note. In: Lingua 12 (1963), 220—225.
Schultink, H.: De morfologische valentie van het ongelede adjectief in modern Nederlands. Van Goor, The Hague 1962.
Trubetzkoy, N. S.: Grundzüge der Phonologie.[6] Vandenhoeck, Göttingen 1977, [1939].

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 135–150 (1986)

OVER THE BORDERLINE: WORDS OR PHRASES?

W. MEIJS

(University of Amsterdam, The Netherlands)

1. Introduction

My topic will be certain types of word-combination in Dutch about whose erivation and status there is disagreement among morphologists. The central uestion is whether such combinations are to be regarded as words or as hrases, and whether their derivation should be accounted for via syntactic or word-formative (lexical) rules. I shall first discuss some characteristics of Dutch synthetic compounds in general, on the basis of the approach adopted in Meijs 1981. I. shall then focus on the particular type or types of syinthetic compounds at issue, reviewing and answering some of the points of crticism raised in Hoekstra 1983, 1984. Finally I shall present some more general considerations concerning the interaction between syntax and word-formation.

2. Dutch synthetic compounds

2.1. Overview

The following features are characteristic of Dutch synthetic compounds:

a) They can be analyzed into two immediate constituents, the second of which is a suffixed derivative.

b) The second constituent gets a fully transparent, compositional semantic reading.

c) The semantic reading of the whole combination is truly synthetic in that the meaning of the first constituent fills a natural gap in the interpretation of the second constituent.

In Meijs 1981 I in fact mentioned as the most striking feature of synthetic compounds that they typically consist of "an existing word followed by a non-existing possible word", or rather of "a listed word followed by an unlisted possible word", where "listed" and "unlisted" are to be taken as the (idealized) competence concepts corresponding to the performance notions "existing" and "non-existing" respectively. Striking though this characteristic may be (it also figures prominently in Roeper & Siegel's 1978 account of verbal synthetic

compounds in English), Van Santen 1986 argues — correctly, I think — that it is probably an effect of the semantic aspects in (b) and (c), rather than an independent basic characteristic of synthetic compounds. It is because they need a further semantic complement that the second-constituent derivatives tend to combine into synthetic compounds rather than occur on their own, and where they do occur on their own this is usually because they have developed lexicalized meanings which include a certain additional semantic element (mostly involving "intensification" of some kind), filling, as it were, the natural gap in their otherwise compositional interpretation. Consider the following combinations from Meijs 1981:

(1) (a)	*wraakneming*	*wraak neem -ing*
	'act of revenge'	'revenge take -ing'
(b)	*huishouding*	*huis houd -ing*
	'household'	'house hold -ing'
(2) (a)	*kortademig*	*kort adem -ig*
	'short of breath'	'short breath -y'
(b)	*hardhandig*	*hard hand -ig*
	'hard-handed'	'hard hand -y'
(3) (a)	*werknemer*	*werk neem -er*
	'employee'	work take -er'
(b)	*notenkraker*	*noten kraak -er*
	'nutcracker'	'nuts crack -er'

The second-constituent derivatives in the (a)-cases (*neming*, *ademig*, *nemer*) do not occur on their own. The ones in the (b)-cases do, but only with idiosyncratic, lexicalized meanings (*houding* 'attitude', *handig* 'handy, skilful', *kraker* 'squatter'). Clearly those lexicalized meanings do not play a role in the interpretation of these synthetic compounds; rather, they obviously get a fully compositional reading, just like the second constituents in the (a)-cases.

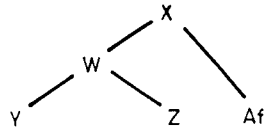
Verbal compounds (as Roeper and Siegel call them — i.e. compounds in which the second constituent is a suffixed derivative of a verbal base) form an interesting subset within the wide range of synthetic compounds. (4) provides a representative sample of the various types that occur:

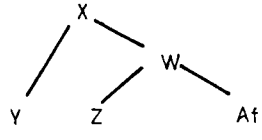
(4)	V-*er*	V-*ing*	V-*end*	*ge*-V-*d*
N	*ijsbreker*	*vormgeving*	*winstgevend*	*handgeknoopt*
	ice break-er	form give-ing	profit give-ing	hand Fp-knot-Fs
	'ice-breaker'	'design'	'profitable'	'hand-woven'
Adj	*witmaker*	*kortsluiting*	*zwaarwegend*	*vetgedrukt*
	white make-er	short close-ing	heavy weigh-ing	bold Fp-print-Fs
	'whitener'	'electric short'	'ponderous'	'bold-printed'

Adv	*langslaper*	*snelpersing*[1]	*langlopend*	*veelgevraagd*
	long sleep-er	fast press-ing	long run-ing	much Fp-ask-Fs
	'longsleeper'	'fast press'	'long-running'	'much-asked'
P	*uitsmijter*	*aanhouding*	*aangrenzend*	*uitgeteld*
	out throw-er	*on keep-ing*	*on border-ing*	*out* Fp-*count*-Fs
	'doorman'	'arrest'	'adjacent'	'K. O. 'd'

(The *Fp* and *Fs* in the English "morphemic transcription" in the fourth column stand for "flexion prefix" and "flexion suffix" respectively, i.e. they are meant to represent the two discontinuous elements (*ge-* and *-d/-t*) that together make up the Dutch past participle morpheme.) The internal make-up of these verbal combinations led me to conclude that synthetic compounds have to be analyzed in terms of the structure suggested for English synthetic compounds in Allen 1978 (cf. 5b below), rather than in terms of that proposed in Botha 1980a, b (Botha 1980c was not available to me at the time of writing) as in (5a):

(5)

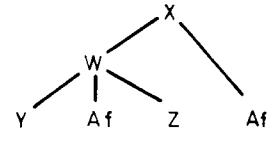


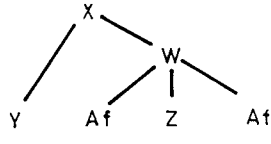


2.2. Past participle combinations

Crucial in the argument is the occurrence and position of the *ge-* prefix of the past participle combinations. Extending the structures of (5a) and (5b) for these would yield the representations given in (6a) and (6b) respectively:

(6)





It is clear that the analysis in (6a) results in a completely counter-intuitive structure, one in which the prefix-part of the discontinuous past participle morpheme becomes some kind of infix within the first constituent. Hence the

[1] I owe this (dictionary) example to Casper de Groot (personal communication) who also mentions *snelberechting* 'fast justice' as another example. I originally thought *Adv* + *V-ing*-combinations constitute a gap within the paradigm of Dutch synthetic verbal compounds, but on the strength of these examples it seems this opinion has to be revised.

"Allen-type" analysis of (6b) seems the only possible one for these combinations. (7a)—(7d) provide some more examples of such past-participle combinations:

(7) (a)	*vetgedrukt*	*vet*	*ge-*	*druk*	*-t*		
	'bold-printed'	bold	Fp	print	Fs		
(b)	*pasgebouwd*	*pas*	*ge-*	*bouw*	*-d*		
	'recently-built'	recently	Fp	build	Fs		
(c)	*goedbeschermd*	*goed*	*be-*	*scherm*	*-d*		
	'well-protected'	well	Dp	protect	Fs		
(d)	*rijkgedocumenteerd*	*rijk*	*ge-*	*document*	*-eer*	*-d*	
	'richly-documented'	richly	Fs	document	Ds	Fs	

Although word-internally adjectives and the corresponding de-adjectival adverbs are not formally distinct, the semantic interpretation of (7c) and (7d) forces an adverbial reading on the first constituents (*goed* and *rijk* respectively), just as in (7b), where *pas-* 'recently, just' — is an unambiguous, non-deadjectival adverb. In (7a) *vet* could be given an adjectival reading, in accordance with he subcategorization of the verb *drukken*: 'print sth bold' rather than 'print sth boldly'. Notice further that in (7c) the derivational prefix (Dp) *be-*, which signals that the verb *beschermen* is denominal, replaces the past participle prefix (Fp), as is the rule with Dutch prefixed verbs (cf. Schultink 1973 and De Rooij 1980) in their past participle form, showing that this verb functions as a morphologically complex verb. The verb *documenteren* in (7d) is also morphologically complex, but here the denominal status is signalled by the derivational suffix (Dp) *-eer*, which does not block the occurrence of past-participle *ge-* Note, finally, that all of the verbs in the combinations in (7) are listed, lexicalized ones.

2.3. "Pseudo" past-participle combinations

Now consider the combinations in (8):

(8) (a)	*breedgeschouderd*	*breed*	*ge*	*schouder*	*-d*		
	'broad-shouldered'	broad(ly)	Fp	shoulder	Fs		
(b)	zwaargebouwd	*zwaar*	*ge-*	*bouw*	*-d*		
	'heavily-built'	heavi(ly)	Fp	build	Fs		
(c)	*hoogbejaard*	*hoog*	*be-*	*jaar*	*-d*		
	'advanced in years'	high(ly)	Dp	year	Fs		
(d)	*rijkgetalenteerd*	*rijk*	*ge-*	*talent*	*-eer*	*-d*	
	'richly-talented'	rich(ly)	Fp	talent	Ds	Fs	

On the face of it, the combinations in (8) are just like the ones in (7). In particular the fact that in (8c) the verbalizing prefix *be-* blocks the occurrence of the past participle prefix *ge-*, just as in (7d), while the verbalizing suffix *-eer* in (8d), as in (7d) does not block it, suggests that the second constituents in (8) are past participles of verbs, just like the ones in (7). There is one important difference, however: while the verbs in the combinations in (7) are all independently occurring, listed verbs; those in (8) are unlisted; i.e. they do not occur on their own as independent verbs with the fully transparent sense involved in these synthetic compounds.

Thus (7b) clearly requires the "normal" listed (existing) verb *bouwen*, meaning 'to build', which is then further modified by the adverbs *pas*, 'recently, just'. In (8b) this cannot be done: a person who is heavily-built is someone with a heavy build, not someone who has been built, constructed, heavily. In other words, the adjective *zwaar* in (8b) is construed with the noun *bouw*, 'build', contained in the transparent verb meaning *voorzien van (een) . . . bouw*: 'provide with (a) . . . build', filling the gap indicated by the dots. Similarly the noun *jaar*, contained in the transparent, unlisted verb *bejaren* has to be construed with the adjective *hoog* in (8c). By contrast, although the noun *scherm*, 'screen' can still be recognized in the verb *beschermen* in (7c) (etymologically something like 'provide with a screen') if one thinks about it, there can be no doubt that in ordinary language use this verb is no longer interpreted decompositionally, so that the interpretation of (7c) is simply 'protected in a good manner' rather than 'provided with a good screen'. A similar contrast holds between (7d) and (8d): in (8d) an interpretation combining the noun *talent* contained in the transparent, unlisted verb *talenteren* with the adjective *rijk*, yielding the interpretation 'provided with rich talents' is possible, whereas in (7d) the interpretation can only be 'richly supported with documents', not 'provided with rich documents'. Finally, as pointed out above, *vet* in (7a) may indeed be given an adjectival reading ('printed bold'), but one that follows the subcategorization of the existing verb *drukken*, '(to) print', rather than along the lines of the cases in (8) ('provided with bold print'). (8a), on the other hand, again shows the familiar pattern of adjective and noun combining in the semantic interpretation: 'provided with broad shoulders' rather than 'shouldered in a broad manner'. Interestingly, there is an existing non-transparent verb *schouderen*, meaning '(to) put on one's shoulders' (said of rifles etc.) which does not allow reference in semantic interpretation to the noun contained in it. Thus *een slechtgeschouderd geweer* does not mean 'a rifle provided with bad shoulders' but 'a badly-shouldered rifle'.

To bring out the difference between "real" and "pseudo" participial combinations three "minimal pairs" are presented in (9—11):

(9) (a) [[[breed]] [ge[[schouder]] d]] (persoon)
Adj Adv N V Adj Adj

(b) [[(slecht)] [ge[(schouder)] d]] (geweer)
adj Adv n V Adj Adj

(10) (a) [[[zwaar]] [ge[[bouw]] d]] (persoon)
Adj Adv N V Adj Adj

(b) [[(slecht)] [ge[(bouw)] d]] (huis)
adj Adv n V Adj Adj

(11) (a) [[[rijk]] [ge[[talent] eer] d]] (persoon)
Adj Adv N V Adj Adj

(b) [[(rijk)] [ge[(document) eer] d]] (geval)
adj Adv n V Adj Adj

In the (b)-cases the internal structure of the listed verbs in the second constituent (marked with round brackets and lower-case letters) is inaccessible to semantic interpretation and the past participle is construed with the adverbial first constituent. In the (a)-cases, on the other hand, semantic interpretation penetrates to the noun within the unlisted verb in the second constituent and construes this "synthetically" with the adjective within the first constituent.

Summarizing, then, morphologically the second constituents in the synthetic combinations are undoubtedly "true", not "pseudo", past participles (albeit of unlisted, non-existing verbs); due to the presence and position of the prefix, only the Allen-type structure will do. The semantic patterning, however, cuts across this, by combining the adjective from the first, and the noun from the second constituent, apparently in accordance with some principle which stipulates that semantic interpretation must go down to the lowest accessible elements. Such combinations thus typically exemplify the kind of "bracketing paradox" discussed in Williams 1981 (or the independence of different levels of representation in Pesetsky's 1985 terms), while at the same time providing further evidence for the scope of the Right-hand Head Rule in Dutch (cf. Trommelen and Zonneveld 1986).

3. Two types of pseudo-participle combination?

3.1. Hoeksema's tests

Hoeksema (1983 and 1984 — the argumentation and the examples are almost identical in both publications) differentiates between two types of pseudo-participle combination. Hoeksema essentially adopts the analysis in Meijs 1981, summarized in the preceding section, but then goes on to argue that combinations like *zwaargebouwd*, 'heavily-built', *slechtgehumeurd*, 'ill-tempered', and *fijngevormd*, 'finely-formed' are phrases rather than words, in spite of the fact that they are normally written as one word. *Breedgeschouderd*, 'broad-shouldered', does have word-status in Hoeksema's view. He is not very clear about what kind of words pattern with *breedgeschouderd*, but I surmise that he would similarly allow word-status to such combinations as *welgemanierd*, 'well-mannered', *goedgemutst*, 'well-cheered', and *hoogbejaard*, 'high-yeared, i.e. advanced in years'. Although I shall argue that Hoeksema's distinction carries no weight, I shall call the two purported types "type I" (the "phrases" *zwaargebouwd* etc.), and "type II" (the "words" *breedgeschouderd* etc.) respectively, for ease of reference. To support the division into two types Hoeksema uses three tests: (a) comparative-formation, (b) fronting of adverb by questioning using *hoe*, 'how', and (c) substitution of the adverb by means of a pro-form such as (*net*) *zo*, 'just so', or *hoe*, 'how'.

As far as test (a) is concerned, Hoeksema argues that the comparative forms of type I are arrived at by adding the comparative suffix *-er* to the first constituent rather than to the whole combination, which yields the following picture:

(12) (a)	*zwaargebouwd*	(b) *zwaarder gebouwd*	(c) **zwaargebouwder*
	'heavily-built'	'more heavily built'	'heavily-built-er'
(13) (a)	*slechtgehumeurd*	(b) *slechter gehumeurd*	(c) **slechtgehumeurder*
	'ill-humoured'	'worse humoured'	'ill-humoured-er'
(14) (a)	*fijngevormd*	(b) *fijner gevormd*	(c) **fijngevormder*
	'finely-formed'	'finer formed'	'finely-formed-er'

For test (b) Hoeksema contrasts the possible constructions on the basis of type I in (15) with the impossibility of similar constructions on the basis of type II as exemplified in (16):

(15) (a) *Hoe zwaar is hij eigenlijk gebouwd?*
How heavily is he actually built?
'How heavily-built is he actually?'
(b) *Hoe slecht is meneer gehumeurd?*
How ill is his lordship humoured?
'How ill-humoured is his lordship?'
(c) *Hoe goed is die stoel gevormd?*
How fine(ly) is that chair formed?
'How finely-formed is that chair?'

(16) (a) **Hoe breed is ie geschouderd?*
How broad is he shouldered?
(How broad-shouldered is he?')
(b) **Al te breed is ie niet geschouderd.*
All too broad is he not shouldered.
('He is not very broad-shouldered.')

To illustrate test (c) Hoeksema points to contrasts such as those between (17) — type I — and (18) — type II:

(17) (a) *Piet is net zo gebouwd als zijn vader.*
Piet is just so built as his father.
'Piet has the same build as his father's.'
(b) *Hoe is de baas vandaag gehumeurd?*
How is the boss today humoured?
'What's the boss's temper like today?

(18) (a) **Piet is net zo geschouderd als zijn vader.*
Piet is just as/so shouldered as his father.
('Piet's shoulders are just like his father's.')
(b) **Hoe is Piet geschouderd?*
How is Piet shouldered?
('What are Piet's shoulders like?')

3.2. Discussion

Notice that the strength of Hoeksema's argument for all three tests rests on the assumption that the various phrasal combinations (*zwaarder gebouwd, slechter gehumeurd* etc.) are indeed derived, via syntactic operations, from the correponding combinations written (and pronounced) as single words (*zwaargebouwd, slechtgehumeurd* etc.), which is then taken as evidence that those single-word forms are themselves in reality also syntactic phrases, in spite of "outward appearance". Hoeksema fails to provide compelling evidence for the correctness of this assumption, however. Since there is nothing to block

the generation of free syntactic combinations like *zwaar gebouwd, slecht gehumeurd* etc., one could simply assume that these are the sources underlying the phrasal combinations.

As far as the individual tests are concerned the following critical remarks are in order:

Test (a): The forms in (12b), (13b), and (14b) are on the whole better than the corresponding (c)-cases, but I do not share Hoeksema's view that the (c)-cases are ungrammatical (nor do most of my informants). Thus a sentence like (19) sounds quite all-right to me:

(19) *Een fijngevormder figuur dan het hare heb ik zelden gezien.*
A fine(ly)-form-ed-er figure than hers have I seldom seen.
'I have seldom seen a finer figure than hers.'

Furthermore, as Hoeksema himself also points out, the comparative is not a very reliably test, since there are various reasons (such as length, stress-pattern, gradability etc.) why comparative-formation by means of *-er* suffixation may not be possible for particular (groups of) adjectives. Thus Hoeksema also rejects comparative forms like *breedgeschouderder* and *breedgerugder* formed on the basis of type II combinations. In other words: the comparative-test is not suitable as a means of discriminating between type I and type II.[2]

The force of test (b) — fronting — is even weaker, in my opinion. After all, by the side of constructions like those in (15) we can also find ones like those in (20), and these are completely parallel to the type-II-based forms in (21), the grammatical counterparts to the forms in (16) which Hoeksema considers to be ungrammatical:

(20) (a) *Hoe zwaargebouwd is hij eigenlijk?*
'How heavily-built is he actually?
(b) *Hoe slechtgehumeurd is meneer?*
'How ill-tempered is his lordship?'
(c) *Hoe goedgevormd is die stoel?*
'How well-formed is that chair?'

(21) (a) *Hoe breedgeschouderd is hij?*
'How broad-shouldered is he?'
(b) *Al te breedgeschouderd is ie niet.*
'All too broad-shouldered he is not.'

[2] Ursula Doleschal drew my attention to the fact that in German, too, there is considerable variation in the way in which comparatives and superlatives of similar combinations are formed, sometimes even leading to double superlatives, as in *meist-bietendster, höchsttalentiert, höchstgestellteste* by the side of *zuspätkommender, weitgereistester* and *hochtalentierteste*.

The contrast between (15) and (20) underlines the general point made above: one need not regard the sentences in (15) as being derived from the lexical items *zwaargebouwd* etc. (for that goes for the sentences in (20), but one can simply see them as being (trans)formed on the basis of the phrases *zwaar gebouwd* etc. Viewed in this way, a more pertinent question to ask is in fact why the forms in (17) should not be as grammatical as those in (16). My answer to this question is that I do not really feel that Hoeksema's acceptability-judgements here are correct. My feeling is that there is not much to choose between (16) and (17) in terms of grammaticality/acceptability. The apparently somewhat higher acceptability of the forms in (16) is probably an optical illusion, resulting from the fact that two of the verbs, viz. *bouwen*, '(to) build', and *vormen* '(to) form) are very frequently used as ordinary, lexicalized verbs, while we should keep in mind that we are here supposed to be dealing with their non-lexicalized transparent counterparts *voorzien van* (*een*) . . . *bouw*/*vorm* etc. ('provide with (a) . . . build/form' etc.), just as the interpretation of *schouderen* is supposed to be the transparent *voorzien van . . . schouders* ('provide with . . . shoulders').

A similar argument holds for test (c) — pro-forms: the possibility of replacing the adverbial first-element can be explained in a perfectly straightforward manner if we assume that forms in (17) are derived from the corresponding phrases *zwaar gebouwd* etc. Assuming an underlying complex lexical item this is not possible, which means that from them we can only form sentences such as those in (22), again completely parallel to the ones based on type II in (23):

(22) (a) *Piet is net zo zwaargebouwd als zijn vader.*
'Piet is just as heavily-built as his father.'
(b) *Hoe goedgehumeurd is de baas vandaag?*
'How good-tempered is the boss today?'

(23) (a) *Piet is net zo breedgeschouderd als zijn vader.*
'Piet is just as broad-schouldered as his father.'
(b) *Hoe breedgeschouderd is Piet?*
'How broad-shouldered is Piet?'

3.3. Further evidence for lexical status

So far for Hoeksema's tests. In addition, there is positive evidence showing that both type I and type II combinations have lexical status. This can be derived from examples such as those in (24) — type I — and (25) — type II:

(24) (a) *Piet's zwaargebouwdheid speelt hem parten.*
Piet's heavily-built-ness is a handicap for him.
'Piet's heavy build is a handicap for him.'
(b) *Gerrit's slechtgehumeurdheid is bijkans spreekwoordelijk.*
Jerry's ill-humoured-ness is well-nigh proverbial.
'Jerry's bad temper is well-nigh proverbial.'
(c) *Fijngevormdheid is een kenmerk van de Jugendstil periode.*
'Finely-formed-ness is a characteristic of the Jugendstil period.
'Elegance is a characteristic of the Jugendstil period.'

(25) (a) *Zijn breedgeschouderdheid maakt hem erg macho.*
His broad-shouldered-ness makes him very macho.
'Being very broad-shouldered makes him very macho.'
(b) *Welgemanierdheid is tegenwoordig weer helemaal "in".*
Well-mannered-ness is nowadays again completely "in".
'Good-manners are nowadays again altogether the "in" thing.'
(c) *Ik vind zijn goedgemutstheid vreselijk irritant.*
I find his well-cheered-ness terribly irritating.
'I find his cheeriness terribly irritating.'

Sentences like those in (24) and (25) show that both type I and type II combinations can be input to further word-formative processing, in this case nominalization, which constitutes further evidence that the distinction into two types seems a spurious one. Moreover, (26) and (27) show that it is impossible to form such nominalizations on the basis of the corresponding phrases, and again that goes for both "types":

(26) (a) **P's zware gebouwdheid . . .* 'P's heavy built-ness . . .'
(b) **G's slechte gehumeurdheid . . .* 'J's bad humoured-ness . . ."
(c) **fijne gevormdheid . . .* 'fine formed-ness . . .'
(27) (a) **zijn brede geschouderdheid . . .* 'his broad shouldered-ness . .
(b) **goede gemanierheid . . .* 'good mannered-ness . . .'
(c) **goede gemutstheid . . .* 'good cheered-ness . . .'

Another test to establish the word-status of both type I and II combinations by reference to further word-formative processing is prefixation. Consider for instance the, admittedly rather contrived, but perfectly interpretable and grammatical forms in (28) and (29):

(28) (a) *pseudo-zwaargebouwd* 'pseudo-heavily-built'
(b) *pseudo-slechtgehumeurd* 'pseudo-ill-humoured'
(c) *pseudo-fijngevormd* 'pseudo-finely-formed'

(29) (a) *pseudo-breedgeschouderd* 'pseudo-broad-shouldered'
(b) *pseudo-goedgemanierd* 'pseudo-well-mannered'
(c) *pseudo-goedgemutst* 'pseudo-well-cheered'

In fact even the combination of prefixation and suffixation on the basis, again, of both "type" I and "type" II forms is possible, as demonstrated in (30) and (31):

(30) (a) *pseudo-zwaargebouwdheid* 'pseudo-heavily-built-ness'
(b) *pseudo-slechtgehumeurdheid* 'pseudo-ill-humoured-ness'
(c) *pseudo-fijngevormheid* 'pseudo-finely-formed-ness'
(31) (a) *pseudo-breedgeschouderdheid* 'pseudo-broad-shouldered-ness'
(b) *pseudo-geodgemanierdheid* 'pseudo-well-mannered-ness'
(c) *pseudo-goedgemutstheid* 'pseudo-well-cheered-ness'

Prefixation by means of the negative prefix *on-*, (*un-*) could in principle provide further evidence for the word-status of type I and II forms. There is a semantic interference factor here, however. Thus in a strictly formal sense combinations such as *onzwaargebouwd* 'un-heavily-built', *onslechtgehumeurd* 'un-ill-humoured', *onbreedgeschouderd* 'un-broad-shouldered' etc. would seem to be O. K., but not semantically. The reason is the fact that negativizing an adjective is often semantically equivalent to forming an antonym, and to form antonyms of most of the forms considered (of both "types"), it is easier to replace the first element by a word expressing the opposite, as shown in (32):

(32) *zwaargebouwd/lichtgebouwd* 'heavily-built/lightly-built'
goedgehumeurd/slechtgehumeurd 'well-humoured/ill-humoured'
breedgeschouderd/smalgeschouderd 'broad-shouldered/slim-shouldered'

However, Dutch linguistic jargon provides us with at least one frequently-used example of *on-*prefixation, as in (33), along with a further suffixed form built on that, as in (34):

(33) *Zin* (88) *is onwelgevormd.*
'Sentence (88) is un-well-formed (i.e. ill-formed).'
(34) *De onwelgevormdheid van zin* (88) *is evident.*
'The un-well-formed-ness (ill-formedness) of (88) is evident.'

In view of Hoeksema's classifying *fijngevormd* and *goedgevormd* as type I cases, I assume that he would also regard *welgevormd* as type I, but again, in spite of the purported phrasal character, the further word-formative processing shown in (33) and (34) contradicts this, pointing to lexical-word status instead.

4. Between syntax and word-formation: borderline cases

4.1. Words or phrases?

The central question which kept cropping up in the preceding section was whether particular combinations should be regarded as words (lexical status) or as phrases (syntactic status). In the discussion I have so far loosely used outward (written) form as a criterion: if a particular combination was written as two or more separate words (e.g. *zwaar gebouwd*) this was implicitly taken as an indication of phrasal status, while a combination written as one word (e.g. *zwaargebouwd*) was considered as proof of (complex) lexical item status. Naturally, the situation is not really quite that simple. If it were, Hoeksema would presumably never even have come up with the suggestion that a form like *zwaargebouwd* could possibly be a phrase. For one thing, it is clear that there are phrases which have lexical status of a kind (idiomatic expressions, for instance), or which can be claimed to be embedded within complex lexical items (in Dutch words like *langeafstandsloper* 'long-distance runner' and *witteboordencriminaliteit* 'white-collar criminality'). For another, the written form is all a secondary representation of the primary spoken form, and we may assume that there is a fair amount of variation in the way in which language-users write combinations of the kind discussed.

Now if written form (as one or two words) cannot be given any decisive criterial value, does that not mean that the argumentation presented in the preceding section is vitiated? I think not. After all, the discussion in connection with (19)—(23) has demonstrated that each of the syntactic operations presented by Hoeksema as evidence for underlying phrasal status can also be applied in such a way as to keep the underlying forms intact as complex lexical items, while additional positive evidence for lexical status has been presented in the discussion surrounding examples (24)—(34). The conclusion has to be, it seems, that we are dealing with borderline-cases: combinations in which (disregarding written form) structures that can be generated both by the syntax and the word-formation component coalesce, and which therefore demonstrate a janus-like character, behaving either as words or as phrases with respect to further processing.

4.2. Interaction between syntax and word-formation

The conclusion suggested above of course raises the question how such combinations must then be accounted for in a theoretical framework. Via syntactic rules, via lexical word-formation rules, or via both? If the latter, does that not then lead to undesirable duplication, with almost identical rules in the syntax and in the word-formation component?

I think the solution must indeed be found in a combination of syntactic and word-formative rules, but not in the sense of duplication, but of interaction between the two kinds of rules. To put this in the right perspective I must elaborate somewhat about the kind of lexicon-model that I take as my starting-point.[3] I assume a full-entry lexicon-model, in which listed (existing) complex lexical items can be retrieved directly, as wholes, from the item-familiar component, just like listed simplex words. This means that (apart from possible function as redudancy-rules relative to listed complex items), word-formation rules basically serve to define the set of all possible novel (unlisted) complex forms. Since the effect of particular word-formation rules is so often illustrated on the basis of existing complex items, most descriptions seem to suggest that the main function of such rules is the creation of lexicalized complex items as permanent additions to the total vocabulary of the language. In practice, however, word-formation rules are more often than not used to meet the communicative demands of the moment, by their capacity to express particular types of relationships between concepts in a concise, schematic way, within the wider sentence-context. Often novel formations are in fact nonce-formations which will never make it to the permanent lexicon, generated (and interpreted) on the spur of the moment, within the sentence- generating process. From this perspective there are only two kinds of things involved: rules and items. The items are all the simplex and complex forms supposed to be contained in the item-familiar lexicon at the assumed fixed point in time for which the model is defined, while the rules are the operations by means of which the given items can be combined with each other to form larger wholes; complex words, phrases, sentences. It is clear that syntactic and word-formative rules must be able to alternate and that there must be interfaces between them: normally the syntactic rules will operate on the output of word-formation rules, but occasionally this may be the other way round. The latter seems to be the case here.

Suppose that application of a number of syntactic rules yields a sequence which can be represented as *X is Adv ge-V-d*, in which *Adv* is the only element (and therefore the head in the sense of X-bar theory) of an Adverbial Phrase. Suppose further that there exists a word-formation rule *A* ↔ *Adv ge-V-d*, then such a WF-rule could use the syntax-generated sequence as input and, as it were, re-interpret it as a complex word of the category *A*. There is no duplication here: via the interplay of independently motivated syntactic rules a sequence is generated which happens to coincide with the internal structure of a particular type of complex adjective, as laid down in the corresponding WF-rule. The syntax does not generate words, but phrases. Only a (presumably) small subset of these are eligible for re-interpretation as words by WF-rules.

[3] In the context of this article I can only give a brief outline of this lexicon-model. See Meijs 1981 for a more detailed discussion.

We must assume that these kind of re-interpretations are subject to severe restrictions, which for the larger part need not be included in the formulation of particular WF-rules, however, as they are in the nature of general constraints on WF-rules as such. This goes for instance for the features generally associated with "lexical integrity": generic interpretation and referential inaccessibility of embedded nouns, the ban on articles, prepositions etc.[4]

5. Conclusion

At first sight the approach suggested here may appear to converge with the one proposed in Botha 1980c, i.e. an essentially syntactic way of deriving (certain types of) synthetic compounds. There are differences, however. Botha's approach stipulates that the affixation-rules operate on syntactic deep structures; in the proposal put forward here they apply to (a restricted number of) syntactically derived structures. Botha's approach has to assume various restrictions on syntactic deep structure, such as the absence of articles, prepositions, plurals etc. In the present proposal no such restrictions are assumed for deep structures: they are simply seen as normal attributes of the WF-component which act as "filters" with respect to the (very limited) possibilities of re-interpreting a syntactically generated word-group as a complex word. Thus if a language user wants to talk about someone as being an *op magistrale wijze gebouwde body-builder* ('a body-builder built in a grandiose way') he can do so; the restrictions which WF-rules are subject to (such as the ban on prepositional adverbials) will then not allow re-interpretation as complex words. However, if the language user talks about a *fors gebouwde body-builder* ('robustly built body-builder') the metamorphosis into complex word (*fors-gebouwde body-builder* — 'robustly-built body-builder') can take place almost imperceptibly.

References

Allen, M. R.: Morphological Investigations. Dissertation, University of Connecticut 1978c

Botha, R. P.: Allen's theory of synthetic compouding: A critical appraisal. In: Alkemade, D. J., van, Feitsma, A., Meijs, W. J., Reenen, P. van and Spa, J. J. (eds): Linguistic Studies Offered to Berthe Siertsema. Rodopi, Amsterdam 1980a, 157–180.

Botha, R. P.: Roeper and Siegel's theory of verbal compouding: A critical appraisal. In: SPIL 4 (1980b), 1–45.

[4] This is particularly clear in another type of phrasal sequence discussed in Meijs 1986, which may be subject to such re-interpretation, viz. *N* + *V-end*, as in *lawaai makend* — "noise making": here the *N* must be interpretable as the head of a NP functioning as object to the verb, and it may not be further modified for lexical re-interpretation to occur (cf. **lawaai-op-de-gang-makend* — 'noise-in-the-corridor-making'). The WF-rule which accounts for the collapse of this type of phrase into a complex word could be formulated as $A \leftrightarrow N$ *V-end*.

Botha, R. P.: Word-based morphology and synthetic compouding. In: SPIL 5 (1980c).
Hoeksema, J.: Zwaarlijvig en breedgeschouderd. In: TABU 13 (1983), 49—70.
Hoeksema, J.: Categorial Morphology. van Denderen, Groningen 1984.
Meijs, W. J.: Synthetische composita: Voer voor morfologen. In: Spektator 10 (1981), 250—291.
Meijs, W. J.: Een regelparadox. GLOT 9—1/2, Lexikon en Syntaxis (1986), 93—104.
Pesetsky, D.: Morphology and logical form. In: Linguistic Inquiry 16 (1985), 193—246.
Roeper, Th. and Siegel, M. A. E.: A lexical transformation for verbal compounds. In: Linguistic Inquiry 9 (1978), 199—260.
Rooij, A., de: Past participle prefix *ge*-deletion and the role of stress in Dutch complex verbs. In: Zonneveld, W., Coetsema, F. van and Robinson, O. W. (eds): Studies in Dutch Phonology. Martinus Nijhoff, The Hague 1980, 159—203.
Santen, A. van: Synthetic compounds: Syntax or semantics? Linguistics 24 (1986), 645—660.
Schultink, H.: Het Prefix *ge-* in Nederlandse (en Duitse) Verleden Deelwoorden. In: De Nieuwe Taalgids 66 (1973), 449—468.
Trommelen, M. and Zonneveld, W.: Dutch morphology: Evidence for the Right-hand Head Rule. In: Linguistic Inquiry 17 (1986), 147—169.
Williams, T.: On the notions "Lexically Related" and "Head of a Word". In: Linguistic Inquiry 12 (1981), 245—274.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 151–179 (1986)

IRREVERSIBLE BINOMIALS IN HUNGARIAN

L. PORDÁNY
(University of Szeged, Hungary)

The aim of the present study is to examine Hungarian fixed-orde conjuncts, or what have variously been called fixed-order coordinates (Abra ham 1950), binomial expressions (Gustafsson 1975), irreversible binomials (Malkiel 1959 and many others), binary pairs or binary compounds, and, especially in the past decade or so, freezes (Cooper and Ross, 1975 and others) — this latter term being not only the most succint one but also the most manageable as it allows us to talk about a freezing phenomenon, and, ultimately and perhaps more importantly, a freezing process or freezing processes. Since, however, the latter notion is not a main concern of this study, the terms freeze(s) and freezing will be used throughout the analysis largely for the sake of brevity.

To make the scope of the investigation as wide as possible, we disregard certain distinctions that are for other purposes made between "irreversible binomials proper" or coordinate compounds like *férj és feleség* 'husband and wife', *igen vagy nem* 'yes or no' or *Rómeó és Júlia* 'Romeo and Juliet'[1] on the one hand, and various "reduplicative word pairs" or reduplicatives and near-reduplicatives, whether they happen to form one word as *csélcsap* 'fickle' or are hyphenated like *icike-picike* 'teeny-weeny', separate like *Csili Csala*, (fictional character), or are in some other way linked as *tit for tat* on the other. We include the group that Malkiel (1959) calls "verbal binomials", like *wheelings and dealings* or *come and go*, as in *easy come, easy go* and will extend the analysis over more complex constructions including trinomials, as e.g. *bort, búzát, békességet* (lit.: 'wine, wheat, peace'), and occasionally, multinomials.

Relatively little attention is paid, however, to the "strength of irreversibility", or, to the fact that the degree of fixedness in word order is variable in certain categories of freezes. In Hungarian as well as in (possibly all) other languages — reduplicatives are absolutely irreversible,

[1] Book titles and titles of poems and other literary works of such and similar binary constructions may become, although not all actually are, freezes.

i.e. "most frozen", while the other end of the scale could be represented by such "non-freezes" or almost non-freezes as English *on and off/off and on* or Hungarian *le-föl/föl-le* 'up and down'. The possible causes of such relative reversibility can partly be deduced from the general rules of freezing that will be explicated, and may partly have a historical background. For the purposes of this paper, however, diachronic considerations have not been entered into the analysis, although it is strongly suspected that in the course of future investigations such factors should be seriously considered in order to eventually arrive at a thorough and complete understanding of the basic and most general processes underlying freezing. Yet at this point we must disregard for instance the interesting-looking—and perhaps potentially significant observation that standard Hungarian *kanadai francia* "Canadian French" has recently shown signs of reversing its order and thus appears to be becoming a mirror, as it were, of English *French Canadian* 'francia kanadai'; and at least equally regrettably we must disregard for instance the story of *Budapest*, an earlier version of which, *Buda-Pest*, was a 19th century convert from the original *Pest-Buda*: any attempt at a thorough investigation of even a few of such cases would go beyond the immediate aims — and the determined length — of this study.

Of the two languages participating in this study, English, or rather general research concerning freezes in English is at an incomparably more advanced state than it is concerning freezes in Hungarian, for which Zoltán Gombocz is one of the very few to have touched upon the question at all, and he did so mainly from an etymological point of view. In fact, English seems to be exceptionally "priviledged" in this respect — as well as in some other respects — given the fact that freezes in English have been studied far more thoroughly than in any other language, with Polish and German being perhaps partial exceptions. (See e.g. Szpyra 1983). It follows then that contrastive studies in freezes are as yet almost non-existent; in fact, even for Polish, Stpyra's above mentioned study is the only contrastive one published so far.

In addition to a generally high level of the description of the English lexicon, there may be at least one more reason why freezes in English rather than in other languages have attracted researchers: as Abraham (1950) and others estimate it, there are a minimum of 500 "binomials proper" (reduplicatives excluded) in English, while Polish and Hungarian, two of the closest runners-up so far, have only about, or not many more than, 250 each (Szpyra's and our estimation, respectively); in other words, English apparently abounds in freezes as compared to other languages. Whether this is actually the case, and if so, whether it is mainly or exlusively due to the highly idiomatic charac-

[2] Throughout this study, quotation marks in parentheses are used to indicate reverse order.

ter of English — idioms, as many authors have pointed out, are "hotbeds of freezing" — is yet to be shown, partly by continuing data-collection in the languages already studied for freezing, and also by drawing into the research languages hitherto not considered.

Considering the apparent predominance of English freezes both in terms of the existing corpus and of the advancement of research, it seemed logical and appropriate at the beginning of this project to use English for a starting point and as a basis. It also seemed useful and logical, then, to primarily concentrate, as almost everyone recently dealing with freezing did, on the sequence of the elements, i.e., to put it very simply, on "... why the ordering of... elements is fixed." (Cooper and Ross, 1975, 63), or, to put it somewhat differently and a little more precisely, what are the constraints or rules that determine the sequence of the elements in freezes.

For a long time it has been known for not only the linguist but also — implicitely — for any intelligent observer that the ordering in freezes is somehow phonologically determined. Cooper and Ross (1975) were the first who claimed to have actually identified and formulated a set of rules that act and interact in freezing, and who also test and demonstrate their findings — in most cases — with ample examples. Thus, one of the basic questions we asked, after much data-gathering and the subsequent construction of a sizable collection of examples in Hungarian, whether and/or to what degree the established rules are applicable to this language, i.e. whether freezes in Hungarian conform to the same rules that govern them in English. One of the ultimate questions of course is that of universality, but since our basic question itself obviously points in this direction, at this point we would not push it any further by making hasty predictions.

Besides a phonological investigation, Cooper and Ross (1975) also extensively deal with the semantics of freezing, given the long-standing observation that the other general factor constraining the sequence of elements in freezing consists of a set of semantic rules.

The basic problems we set forth to examine can thus be summed up as follows:

1. The applicability to Hungarian, and possible interaction of, phonological rules established for English;

2. The validity of English, and purportedly some general, semantic constraints;

3. The (possible) interaction of phonological and semantic constraints; and finally,

4. Prospects of universality.

*

For phonological studies, the circle of freezes has to be practically restricted to reduplicative compounds, since minimal pairs (or: near-minimal pairs) necessary for the identification of regularities are, apart from a few exceptions,[3] not to be found among "binomials proper", i.e. coordinate phrases like *father and son, here and there, sooner or later*,[4] and so forth. Reduplicatives, however, do by nature exhibit precisely the characteristics that make phonological considerations possible.[5]

Cooper and Ross (1975) identify altogether seven phonological rules that, according to the authors, determine the sequence of the elements in freezes.[6] Originally, all seven were to be thoroughly applied to Hungarian, however, it soon became clear that, for various reasons to be indicated below, at least three out of the seven had to be dropped out of consideration at the beginning. The excluded rules and the reasons of their exclusion are as follows:

1. Longer resonant nuclei, according to which the second element of the freeze contains longer resonant nuclei than the first, other factors being equal.

This is quite obviously the shakiest and the most tentative of the suggested principles. The total number of examples the authors found is a exceptionally and discouragingly low t w o, and neither is a clear minimal pair at that. On top of this, the "longer resonant nucleus" of the second element, even if better supported, would at least partly be based on quantitative differences between monophthongs and diphthongs,[7] which would make a comparison hazardous as the Hungarian vowel system is not characterized by dipthongization. All this should not necessarily exclude an examination of Hungarian in respect to the suggested rule in question; what really made us

[3] Exceptions include mainly 1) cases where semantic constraints do not prevent the freeze from conforming to the powerful s y l l a b l e l a w (see below), and 2) coordinates that are at the same time "almost reduplicatives", as e.g. H. *papás-mamás, hetet-havat* etc.; such examples are freely included in illustrative lists wherever possible or necessary.

[4] As it will be shown, the vast majority of such freezes are also ill-suited for phonological considerations because their ordering is mostly determined by semantic factors.

[5] The renderings into English in general, and especially those of reduplicatives are throughout this study often nothing more than rough estimates. It should be kept in mind that many reduplicatives are nonsensical or near-nonsensical (in both English and in Hungarian), while at the same time they often carry some sort of "flavor" and many of them are onomatopoeic or onomatopoetic. All these things make translation difficult. A parenthetical explanation, rather than a gloss or translation is attempted wherever the latter seemed too difficult or impossible.

[6] Some of these rules have inces undergone a number of various modifications as seen e.g. in Oden and Lopes (1981) and Szpyra (1983), none of which resulted, however, in any significant change, and practically none effecting the present comparison. It seems therefore sufficient to take the original set of rules as a starting point.

[7] One of the two examples is s t r e s s a n d s t r a i n, where the diphthongal nature of the nucleus in the second element is obvious, and in place 2 element of the second example, t r i c k o r t r e a t, there frequently is at least some diphthongization.

neglect it for the time being was the fact that the Hungarian corpus showed no minimal pairs in support of it.

2. Decreasing obstruency, according to which, if both elements end in a single consonant, then the second element has a less obstruent final segment.

This principle is much better supported than the former: there are over a dozen pairs (minimal or near-minimal) listed for English, including *push and pull, rock and roll, hit or miss* etc. The problem with comparison is that Hungarian seems to lack — apparently entirely — this particular constraint. There are a great number of minimal and near-minimal pairs where both elements end in a single consonant, as the following random list exemplifies:

iru*l*-piru*l*	'keep blushing'
gilin*g*-galan*g* bi*m*-ba*m*	'ding-dong'
ákom-bákom	'scrawl, scribble'
ille*g*-bille*g*	'have a rolling gait'
dino*m*-dáno*m*	'merrymaking'
csi*p*-csu*p*	'petty, trivial'
szánom-bánom	'repentance'
te*sz*-ve*sz*	'potter about'
zsong-bong	'buzz, drone, ring'
csapo*t*-papo*t*, as in *otthagyott csapot-papot*	'abandon, leave suddenly, take French leave'
vagy megszoki*k* vagy megszöki*k*	(this will either) 'make you or break you'
tipe*g*-topo*g*	'toddie about, patter about'
eszem-iszom	'feasting' etc.

The list could be extended to well over fifty items and possibly to nearly a hundred. However, as it is demonstrated by the examples, the final consonants are always the same in the two elements, as if they were to conform to some sort of a rhyming principle, so much so that up to now we have found no counterexamples to this pattern.[8] Thus, we may at least tentatively state the principle that in Hungarian minimal (and near-minimal) pairs, the final consonants of the elements are identical, a rule that applies not only qualitatively but also quantitatively, as shown for instance by the following examples with "long" or "double" consonants in word-final positions concerning both elements:

[8] *Ipi-apacs* (a hide-and-seek game) looks like the one counterexample found however, it is clearly a variation (apparently for easier pronunciation) of original and semantically identical *ipi c s-apa c s.*

ütött-kopott	'shabby, threadbare'
rissz-rossz	'worn out', 'cast off'
itt-ott	'here and there'
csipp-csöpp or csipp-csepp	(sound of raindrops falling)
kipp-kopp	'tap-tap'
piff-puff, dirr-durr	(sounds of shooting, or slapping in the face)
szedett-vedett	'trashy, shoddy, scrappy'
lics-locs etc.	(sound of water splashing)

The strength of this principle seems to be supported by the fact that it is also applicable to many (probably a majority of) non-minimal and often semantically determined pairs, although it is clear that in most of these cases the identicality of the final consonants is determined by suffixation rules in Hungarian:

foggal-körömmel	'tooth and nail'
tűzzel-vassal	'by fire and sword'
térül-fordul	'go and get quickly'
csetlik-botlik	'stumble about'

3. Fewer final consonants, according to which, other factors being equal, the second element of the freeze should contain fewer final consonants than the first.

The inapplicability of this principle to Hungarian is clearly demonstrated above, to which we might add that it is a highly tentative one to begin with. The authors were unable to list any clear pairs in support of it, and the number of their near-minimal pairs is a very low three, to which, in addition, they immediately add two counterexamples. Thus, the best we can do at this point is echo the authors who remark that they are not sure whether this principle should (altogether) be given up (Cooper and Ross 1975, 77).

With three of the proposed rules eliminated and the scope of investigation narrowed down, our attempt at a thorough examination of Hungarian is, luckily, still not in as much danger as it might look at first, in spite of the fact that there is a forth rule that presents some problems of both validity and usability for comparison (see further below). That this is the case is suggested — in addition to the fallibility and/or relative insignificance of the first three rules as already demonstrated — by the hierarchical o r d e r i n g of the rules. When Cooper and Ross tried to examine the relative strengths of their seven phonological rules, i.e. when they tried to establish or approximate a ranking order between them based on their respective overriding strengths, two out of the three in question fell in last and next to last places on a scale of a decreas-

ing order of strength. And although the remaining third rule ("longer resonant nuclei") is put in second place, we see it as practically unsupported and hence at least highly doubtful, and can but consider it one of those " . . . subparts of this hierarchy . . . [which are] by no means firmly established, . . . " (Ibid., 80).

Let us then briefly turn to the fourth rule, which we referred to above as problematic. This rule concerns the o b s t r u e n c y of initial consonants and states in its original form that " . . . the obstruency of the initial consonant of a place 2 element will be stronger than the obstruency of the initial consonant of a place 1 element, other factors being equated." (Ibid., 75).

The increasing order of obstruency as presented by the authors is the following:

glides → liquids → nasals → spirants → stops.

The examples in English are numerous:

wear and tear
walkie-talkie
roly-poly
hickory dickory
razzle-dazzle

are just a few of the several dozen found by the authors.

The problem, however, starts with the fact that there are at least two groups of freezes that constitute exceptions to the proposed rule. One of the two, represented by examples like

tootsy-wootsy
teeny-weeny
pall-mall etc.,

brings into the picture yet another rule having long been observed by many researchers and being called l a b i a l i z a t i o n, or more exactly, p l a c e 2 i n i t i a l l a b i a l i z a t i o n, which Cooper and Ross do not include in their set of rules but are aware of it and also of the fact that it is in conflict with i n c r e a s i n g o b s t r u e n c y or rather, that their principle of obstruency is in conflict with labialization. This is in fact the main reason why they are uncertain about their own proposition, which is quite clear from a footnote remark: "We have chosen to argue for an alternative obstruency-based account not because we are convinced that it is right, but because we hope that future researchers will be able to find crucial evidence that will resolve our present dilemma." (Ibid., 108).

The existence, and marked effect of labialization in English is mentioned by Jespersen (1961); it is shown by several authors (e.g. Jacobson 1972, and

Szpyra 1983)[9] to exist in Slavic languages, and although it is no intention of ours to try to solve the authors' dilemma at this point, we might note that labialization is also markedly present and strongly operative in Hungarian in several ways. Of the various labials, this language seems to have a strongly favored one, namely the voiced bilabial stop, which "labializes" conjuncts through either 1) replacing the place 1 initial consonant in the place 2 element, or 2) appearing additionally as the initial consonant of the place 2 element, as shown by the following lists of examples:

1) *cs*iri-*b*iri	'trifling'
*cs*ere-*b*ere	'swapping'
viminek a *cs*inja-*b*inja	'the ins and outs of sg, the hang of sg'
*cs*onka-*b*onka	'crippled'; 'incomplete'
*cs*ecse-*b*ecse	'knick-knacks, knick-knackery'
*s*úg-*b*úg	'susurrate'
*r*ontom-*b*ontom	(scolding rhyme)
*z*úg-*b*úg, *zs*ong-*b*ong	'rumble/ring/toll with a deep tone'
*h*elló-*b*elló, *sz*ia-*b*ia	(playful variations of 'hi, hello')
*t*ündi-*b*ündi	'leary'
2) ákom-*b*ákom	'scrawl, scribble'
ingó-*b*ingó	'unsteady, wobbly'
ingyom-*b*ingyom, egyedem-*b*egyedem adta-*b*adta(-tereintette)	(bits of nursery rhymes)
encsem-*b*encsem	'tidbit', 'kickshaw'
ugra-*b*ugra, (ugra-*b*ugrál)	'skip about'[10]

This labial seems to have the function of a general device — especially in children's language as for instance in nursery rhymes — of playful reduplication, since it can be glued, as it were, onto the front end of a number of nouns, thus forming a nonsensical reduplicative of a basically meaningful word:

uram-*b*uram	('my lord' — *b* + duplication)
erdő-*b*erdő	('forest' — *b* + duplication)
ejnye-*b*ejnye	(expression of mild scolding)

[9] Szpyra, who lists a number of examples from Polish, where place 2 initial segments include bilabial nasals and bilabial stops, asserts that "The evidence for this rule in Polish is overwhelming . . ." (Szpyra 1983, 46).

[10] The "b-effect" is not exclusive. Both types 1) and 2) can also be observed, although much less frequently, as occurring with the voiceless labial stop, as in *locs-pocs* 'sludge', 'slushy weather' or *ec-pec* (counting out rhyme), respectively. Type 1), in addition, has a few occurrences with the voiceless labio-dental stop, as e.g. in *laca-faca* (*nem laca-facázik* — approx.: 'no funny business; take the matter seriously'), and *locsi-fecsi* 'gossipy', 'chatterer'.

As yet another instance of the powerful "b-effect", it crops up in "incorrect" (sub-standard) Hungarian in place of a Standard Hungarian labio-dental fricative, as in

szer*b*usz for szervusz 'hi, hello'
Zsi*b*ágó for Zsivágó 'Zhivago'
taku-*b*aku for taku-vaku 'trifles' etc.

However, labialization in general and the "b-effect" in particular do not bring us much closer to a solution even despite *roly-poly*, *pall-mall* and *wham-bam*, respectively: in English the problem is further complicated by the second of the two above-mentioned problematic groups of freezes, which include, among others,

boogie-woogie
bigwig and
bow-wow,

whereby we have on our hands a labial as the initial sound of place 1 elements, and not just any labial at that, but precisely the one that is so heavily marked in Hungarian as an exclusively place 2 initial. On the other hand, there is no contrast with either Hungarian or other English groups to talk about either, as the second element in each of the above examples also starts with a bilabial.

The precariousness, or shall we say, messiness of the situation, aggrevated by the authors' own uncertainty about it, makes it advisable that we keep clear of it at least for the purposes of the present study. And since our purpose is to apply to Hungarian rules well-established (or more or less so) for English, we have to altogether exclude not only the obstruency-labialization dilemma, but also the decreasing obstruency-principle itself, from further examination.

Thus we are left with the core of Cooper and Ross' phonological findings, i.e. with rules that are a) well-supported, b) highly-placed in the hierarchy of strength and c) considered convincing by the authors themselves. It is in terms of these principles that a more thorough investigation of Hungarian freezes seemed well-based and justified and that, we believe, it was also possible to gain new insight concerning at least one of the rules, as it will be demonstrated later.

The three rules in question can be stated as follows:[11]

Rule No. 1., the "**syllable law**" (also called **Panini's law**)[12] says that, other factors being (more or less) equal, the number of syllables in the second element

[11] For rules 1 and 2, we add "weaker versions", i.e. slightly extended versions.
[12] So called because it was alledgedly first observed and developed by Panini.

exceeds that in the first, or, to put it in a "weaker version", the first element should not contain more syllables than the second. Somewhat logically, this rule is also often referred to as the s h o r t+l o n g rule.

According to **Rule No. 2.**, or the **consonant rule,** the second element, other factors being equal, contains more initial consonants than the first,

and **Rule No. 3.**, also called **F2,** says that the second element in the freeze contains a vowel with, to use acoustics phonetics terminology, a lower second formant frequency, lower, that is, than for the vowel in the first element. This means in practice that the sequence of vowels for American English should be the following:

i > I > ɛ > æ > a (hot) > ɔ (hall) > u

which, ↓ in a loose approximation[13] can be said as corresponding to Hungarian

i, í (ü) > (ë), (ö) e > á > a (o) > u

The rules having been applied to a relatively large Hungarian corpus — a total of well over a hundred word pairs — the following picture emerges:

1. For Panini's law (the syllable law):

As we pointed out before, there seems to be considerable support for this principle (as well as the other two) in English, despite the fact that it is only applicable to cases with place 1 monosyllabic elements i.e. to freezes that show a 1 + 2 syllabic pattern, as for instance in

hot and heavy
bread and butter
free and easy
bread and water
rough and ready, and so forth.

In Hungarian, while it has a number of examples of the same syllabic pattern, as e.g.

fúr-farag	'keep busy at woodwork'
csűr-csavar	'pettifog'
most vagy soha	'now or never'
rúg-kapál	'kick about'
Góg és Magóg, Vér és arany	(often quoted poem titles)
fel s alá	'up and dow '
etc.,	

[13] An *approximation* is not only quite sufficient for this comparison but is at the same time the only possibility, the two systems of vowels being significantly different in terms of both quality and quantity.

the manifestations of the syllable law are not restricted to the 1 + 2 pattern at all. Firstly, there are examples of a 1 + 3 pattern, as in the following:

kín-keserves	'slow and painful'
bús-borongós	'gloomy, melancholic'
bús-keserű	'heart-sore'
férj és feleség	'husband and wife'
pénzt vagy életet	'money or life'
etc.,	

the latter two being also semantically determined.

Further, there is a frequent 2 + 3 pattern; e.g.:

csűrés-csavarás	'pettifogging', 'turns and twists'
béke és barátság	'peace and friendship'
írni-olvasni	'read and write' ("write and read")
sírás-nevetés	'cry and laugh'
napok és hónapok	'days and months'
szívvel-lélekkel	'with heart and soul'
foggal-körömmel	'tooth and nail'
síró-picsogó	(dialectal nursery rhyme) etc.,

and even a 2 + 4 pattern, for which we found two examples:

híres, nevezetes	'famous, renowned'
adta, teremtette	(scolding rhyme)

There is a very large group, consisting mostly but not exclusively, of reduplicatives (minimal os near-minimal pairs), that shows a 2 + 2 pattern, as

árkon-bokron (túl)	'(over) hedge and ditch'
keszekusza	'dishevelled', 'disorderly'
setesuta	'clumsy', 'awkward'
ihog-vihog	'giggle, snigger'
illeg-billeg	'have a rolling gait'
ázik-fázik	'shiver in the rain and cold'
izeg-mozog	'fidget', 'be restless'
szedett-vedett	'trashy, shoddy'
hebeg-habog	'hem and haw'
hébe-hóba	'now and then'
dínom-dánom	'merrymaking'
nyifeg-nyafog	'keep whining'

ütött-kopott	'shabby, threadbare'
üggyel-bajjal	'with great difficulty'
térül-fordul,	'go and get quickly,'
and so on;	

and there is an almost equally large group showing a 1 + 1 pattern:

lics-locs	'sound of water splashing'
lim-lom	'junk'
ripsz-ropsz	'snip-snap'
rissz-rossz	'worn out', 'cast off'
le-föl	("down and up")
bim-bam	'ding-dong'
üt-vág, üt-ver	'thrash'
itt-ott	'here and there'
ad-vesz	("sell and buy")
óg-móg	'grumble'
ki-be etc.,	("out and in")

The syllable law can be extended to trinomials and multinomials, an area where it seems to operate with an even greater force, conforming to the "strong version" in almost every single case. The examples we found — and they are quite numerous in Hungarian — invariably show a partial or gradual increase in the number of syllables:

zsip-zsup, kenderzsup	(children's rhyme)
bort, búzát, békességet	(whine, wheat and peace)
így-úgy-amúgy	'this way and that', 'anyhow'
itt, ott, amott	'here, there, everywhere'
adta-badta-teremtette, megy, megy, mendegél, ásó, kapa, nagyharang csip-csip csóka	(bits os nursery rhymes)
huj, huj, hajrá	(utterance of rooting/cheering)
bor, sör (or *sör, bor*), *pálinka*	'wine, beer' (or: 'beer, wine'), brandy'

— a common sign in restaurants or bars, with the order of the first two elements being interchangeable but with the third one, *pálinka*, always in third place, *atya, fiú, szentlélek* 'Father, Son and the Holy Ghost', this latter one being (also) semantically constrained.

One partial counterexample we found is *jelen, múlt jövő* 'present, past, future', but this is chronological ordering, a type of semantic constraint that will be touched upon later.

The last three days of the Hungarian week are

péntek, szombat, vasárnap, 'Friday, Saturday, Sunday',

which, in terms of syllable pattern, corresponds precisely with more usual English

'Thursday, Friday, Saturday'.

The four swimming strokes are:

mell, hát, gyors, pillangó ('breast, back, free, butterfly' — in that order), and although there is some disagreement among native speakers as to the ordering of the first three strokes (some say there is no definitive ordering) they almost invariably put *pillangó* in the fourth place. (Competitive swimmers, who use the expression more frequently, consider *pillangó* as absolutely frozen in 4th place.)

Finally, it is perhaps worth noting about Panini's law that, apparently, it also has an effect on a great number of actual occurrences of what we can call non-freezes, i.e. r e v e r s i b l e b i n o m i a l s. Here is one representative example: Ady, the poet has some twenty poems titled by (reversible) binomials linked by *és* 'and', out of which eleven, or over 50% confirm the strong version of the syllable law (in other words they show the syllable patterns of 1 + 2, 2 + 3, 1 + 3 or 3 + 4); another three conform to the weak version (1 + 1, 2 + 2, 3 + 3); and it is only the remaining six, or 33%, that go against the rule (2 + 1, 3 + 2 or 3 + 1) about a half of which are obviously semantically constrained, as e.g. *Az anyám és én* — 'My mother and I'. The titles examined are listed below, in alphabetical order.

Anya és leánya	'Mother and her Daughter'
Asszony és temető	'Woman and graveyard'
Az anyám és én	'My mother and I'
Góg és Magóg	'Góg and Magóg'
Gyűlölet és harc	'Hatred and Fight'
Halottan és idegenen	'Dead and a Stranger'
Harc és halál	'Struggle and Death'
Harcos és harc	'Fighter and Fight'
Kín és dac	'Pain and Spite'
Margita és Ottokár	'Margita and Ottokár'
Margita és sorsunk	'Margita and our Destiny'
Mária és Veronika	'Mária and Veronika'
(Örök) harc és nász	'(Eternal) Fight and Love'
Pénz és karnevál	'Money and Carnival'

Ruth és Delila	'Ruth and Delila'
Seregély és galamb	'Starling and Dove'
Szerelem és ravatal	'Love and laying in state' (lit.: 'Love and Bier')
Vér és arany	'Blood and Gold'
Vihar és fa	'Storm and Tree'
Vulkánok és szívek	'Volcanoes and Hearts'

All in all, it seems that Hungarian can safely be added to the list of languages that clearly confirm Panini's syllable law as is clear from the above; in addition, in Hungarian we do not to worry about cases like

yippity-yap
hippity-hop
flickety-flack
(*hickory-*) *dickory dock*
blankety-blank or
clackety-clack,

all examples of something like a reversed Panini's law (possibly constrained by English rhythmic patterns — see Cooper and Ross (1975, 78)).

The counterexamples we have found for Hungarian are very low in number, and they are generally governed by chronological or other forceful semantic factors, as

kezdet és vég	'beginning and end', or
észak-dél	'North-South'.

One counterexample for which there is no apparent explanation — phonological or semantic — is

háború és béke	'war and peace';

all we can do for the moment is put the blame on Tolstoy,
(cp. Russian *Война и Мир* 'War and peace').

Rule No. 2., the consonant rule, like Panini's law, can be rephrased into a weaker version by saying that the number of initial consonants in place 1 elements should not exceed that in place 2 elements, which at the same time allows for an infinitely larger number in place 2 elements, with special regard to the fact that there is a large group of freezes in Hungarian in which the first element begins with a vowel. This goes especially for many reduplicatives and

near-reduplicatives including examples of the "b-effect" discussed earlier, as for nstance:

adta-badta (*-teremtette*)	(nursery rhyme)
ágas-bogas	'branchy'
ákom-bákom	'scrawl, scribble'
ámul-bámul	'gaze'
árkon-bokron (*túl*)	(over) 'hedge and ditch'
áta-bota (in: *áta-botában*)	'slap-dash'
illeg-billeg	'have a rolling gait'
inog-binog	'move unsteadily'
ugra-bugrál	'hop about'
ügyes-bajos etc.,	'troublesome';

other "labialized" pairs as e.g.:

ad-vesz	"sell and buy"
agyba-főbe (*ver*)	'rough up'
ázik-fázik	'shiver from the rain and cold'
ici-pici	'tiny'
icike-picike	'teeny-weeny'
ihog-vihog	'giggle, snigger'
irul-pirul	'keep blushing'
izeg-mozog	'fidget, be restless'
ízig-vérig	'out and out', 'to the core'
óg-móg	'grumble'
össze-vissza etc.,	'topsy-turvy';

as well as non-labial ones:

ütött-kopott	'shabby-looking'
épülő-szépülő	'be getting more beautiful'
apraja-nagyja	("old and young")
ötöl-hatol	'hedge, beat around the bush'
ipam-napam	(dialectal terms of kinship)
apa, cuka, . . .	(nursery rhyme)
életre-halálra	'for life and death'
élve vagy halva	('alive or dead')
ihaj-csuhaj	(exclamation of boisterous joy)
éjjel-nappal	("night and day")

A second group is represented by zero initial consonant in both elements, as in:

erre-arra	'this way, that way'
itt-ott	'here and there'
eszem-iszom	'feasting'
emide-amoda *ide-oda*	(to) 'here and there'
immel-ámmal	'reluctantly'
innen-onnan	(from) 'here and there'
így-úgy, így vagy úgy, *emígy-amúgy*	(lit.: 'in this way or in that way');

and there is a third group where both elements begin with a single consonant:

csip-csup	'petty, trivial'
hepe-hupa, (hepe-hupás)	'uneven, bumpy'
térül-fordul	'go and get quickly'
szánom-bánom	'repentance'
(összehord) hetet-havat	'talk nonsense'
(otthagy) csapot-papot	'abandon, leave suddenly'
tűzzel-vassal	'by fire and sword'
dínom-dánom	'merrymaking'
cica-mica	'pussy' (term of endearment)
handa-banda	(loud) 'bragging'
sete-suta	'clumsy', 'awkward'
tör-zúz	'smash violently'
hebeg-habog etc.,	'hem and haw',

and many more.

Of the one hundred or so examples examined, we have found no counterexamples of the weak version, i.e. we are aware of no cases with a higher number of initial consonants in place 1 elements. Similarly to the case of Rule No. 1, Hungarian supports the rule more clearly than English, this time not because of English counterexamples, but because of a significantly higher number of supportive examples in Hungarian: in comparison with the 2 clear minimal pairs found for English by Cooper and Ross,[14] our lists contain nearly a dozen for the strong version only, and these lists are far from being exhaustive.

To sum up their joint effect, it seems that Rules 1 and 2 form an alliance, as it were, to reduce the first element of the freeze as much as possible.

[14] *Sea and ski*, and *money shmoney* — the second one being actually Yiddish.

Rule No. 3., the vowel rule, looked equally promising at first sight. In fact, it is far the most extensively supported of all the phonological rules worked out for English, and the Hungarian examples seemed to validate the rule just as convincingly as they did for the other two rules. Both languages abound in examples showing the sequences i (I, í), >e, H ü >œ a(á) o(ó), with hardly a trace of the reversed order under any circumstances.

In addition, Cooper and Ross identified at least five or six of vowel sequences (there are some unclear cases) for which there are minimal pairs. As it is shown by the examples below, each of these sequences is applicable and valid for the Hungarian corpus; therefore we have a number of almost one-to-one correspondences for English and Hungarian (with slight differences of vowel qualities in the two languages disregarded):

	Am. English	Hungarian
[I(i) > ɔ:(o)]	*ping pong*	*itt-ott*
	ding dong	*rissz-rossz*
	singsong	*lig-lóg*
	crisscross	*ripeg-ropog* etc.
[i > u]	no 2nd place minimal pairs, but:	*úgy vagy úgy*
		dirr-durr
	bibbity bobbity boo	*csip-csup* etc.
AmE [I > a]	*tick tock*	*dínom-dánom*
	flip flop	
[e > a]	*by guess and by gosh*	*hetet-havat*
		hebeg-habog etc.[15]

There is one particular sequence: [I > œ], for which alone there are twelwe minimal pairs (*mishmash, fiddle-faddle, chitchat* etc.) but for which we have found no similar pairs in Hungarian. However, this is still no problem as there are non-minimal pairs for the same sequence including

	illa-berek	(nursery rhyme)
and	*ki-be*	("out and in")

with *enni-inni* 'eat and drink' being the only counterexample found, which is clearly semantically determined — something to be discussed separately.

The problem starts with our initial observation that in Hungarian, the F2 principle holds only as long as there is a high front vowel in the

[15] These entries are not semantically comparable to those in the left-hand column. They are glossed elsewhere in this study.

first element and whatever the second contains is a low and/or back vowel with relation to it, as can actually be seen in all the sequences shown above. In Hungarian, in other words, the F2 rule does not work for pairs with a back vowel in their first elements. Especially striking in this respect is the frequent [u → . . .] pattern. The [u] sound, with the lowest second formant frequency of all the vowels, should not occur in place 1 elements to begin with, or at best it should be restricted to reduplicatives and near-reduplicatives as for instance

pooh-pooh or
hook and crook (as in: *by hook and crook*) or
choo-choo (as in: *choo-choo train*).

This, however, is far from being the case in Hungarian. In fact, the sequences [(o) > a > u] proposed by Cooper and Ross occur precisely in the reverse in an overwhelming majority of the cases, as e.g. in:

fúr-farag	("keep busy at woodwork")
húz-von	'keep putting off'
kutya-macska	('dogs and cats')
csúszik-mászik	'creep, grovel'
rúg-kapál	'kick about'
búbánat	'sorrow, grief'
bús-borongós	'gloomy, melancholic'
huj, huj, hajrá	(expression of cheering/rooting)

while the only example that we have found clearly confirming F2 is

lót-fut

and this one may well go back to historical reasons.
What helped solve the problem is what looked most discouraging at first. Namely, upon closer examination it turns out that the F2 principle does not really work for English, either. What first aroused suspicion was that the proposed vowel sequences for which Cooper and Ross did not find minimal pairs include [a > u] and [o > u], for the exact reverse of which we have a number of examples (although not minimal pairs) in Hungarian, as can be seen above. Then the authors remark at one point that they have found "one serious counterexample" to their proposed ordering in English

ooh and aah.

Now, to this we can easily add the trinomial-like interjection
brou-ha-ha (*brouhaha*), and, upon some further search, a number of other pairs with nearly the same pattern that have escaped the authors'

attention, as e.g.

foot and mouth (disease)
hook and eye
room and board
root and branch
bull and oyster (this latter one has the syllable rule going for it, too).

And should these examples not by themselves arouse enough suspicion, then, for good measure and some further cross-linguistic control testing, we can consider the following examples from German:

(*von*) *Ruf und Rang*
(*Das*) *Drum und Dran*
Lust und Laune
Sturm und Drang
(*Der Ritter ohne*) *Furcht und Tadel*
(*über*) *kurz und lang*, etc.,

with *Hab und Gut* being the only counterexample that comes to mind without a more extensive search.

The above examples in three languages are in themselves so overwhelming that one can hardly resist the temptation to add to the list — going back to English again:

Cooper and Ross (u > a)

notwithstanding the syllable rule or alphabeticism, respectively, or whatever other reason for this ordering.

But guessing aside, the point of course is that what we are dealing with is not just a large group, or groups, of random exceptions from different languages from an otherwise and basically valid rule, but that a new rule is emerging, or at least the F2 has to be considerably modified.

Considering the lists of examples confirming the F2 rule for both English and Hungarian, and also the above "exceptions" in three languages, the new rule can be shown in a simple way. As we have seen, the *u*, a high back vowel, is almost never followed by a high front or even a low front one; the sequence is almost exclusively from high back to lower back.

In the cases where the F2 rule does seem to work perfectly (i.e. with the high front vowels), a similar high to low and/or front to back shift in the points of vowel formation can be observed. Consequently, the F2 could — and we argue that it should — be replaced by a lower and back, or "low-back" rule, covering both higher to lower and/or front to back movements.

The vowel charts below (Figure 1 and Figure 2) are a diagrammatical illustration of the original rule of the modified one, respectively.

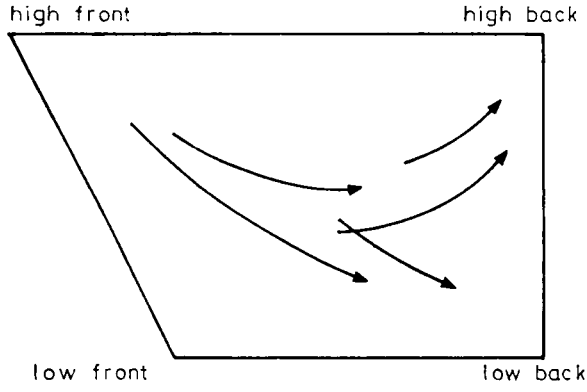


Fig. 1

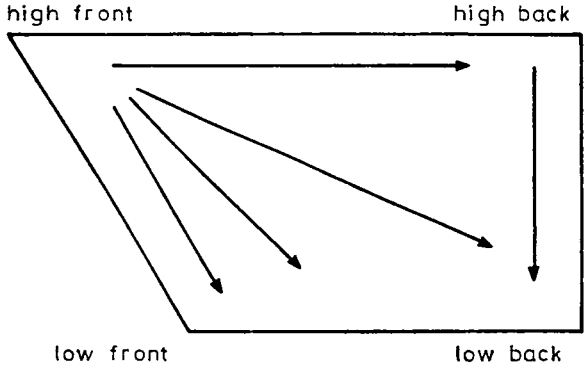


Fig. 2

Some question marks remain; consider e.g. English

calm and cool, as especially in:
calm, cool and collected; or Hungarian

jár-kel — we shall come back to this one. But the number of counterexamples is, for one thing, incomparably smaller than that for the originally proposed rule. All in all, both English and Hungarian conform to a single vowel rule with about the same accuracy that we saw for the other two rules discussed.

We may also note in passing that, according to our preliminary observations, the vowel rule (as well as perhaps the other two rules) may somehow

operate not only in simple freezes but also in longer units including proverbs and similar constructions, in a fashion observed in freezes, and not entirely differently from what Cooper and Ross suggest for English proverbs in s e m a n t i c terms.[16]

For a tentative and random illustration of what might be emerging, consider the following Hungarian sayings:

1. *Amit n y e r a r é v e n, elveszti a v á m o n* (roughly: 'What is lost on the swings is made up on the rounds' — order of clauses reversed), and

2. *Jobb ma egy v e r é b, mint holnap egy t ú z o k* (roughly: 'A bird in the hand is worth two in the bush').

What is worth noting is apparently the syllable pattern and the vowel sequences for place 1 and place 2 elements.

If we want to provide a summary of our phonological findings so far and if we want to do it in terms of the question of possible u n i v e r s a l i t y, we have good reason to be a little more optimistic than were Cooper and Ross, who ventured to go only as far as saying that " . . . The prospects of universality for c e r t a i n phonological constraints on freezing are s o m e w h a t brighter" i.e. than for semantic constraints.[17] As we have seen, all well supported rules, whether originally worked out for English or for Hungarian, apply to both languages with more or less equal strength and show a more or less equally low number of exceptions in both languages.

The one question we cannot even attempt to give an approximate answer to at this point is that concerning the relative strengts of a rule over another, i.e. possible overriding effects among rules; in other words, we do not, or not yet, seem to have enough evidence to seriously suggest any hierarchy or ranking of strength. The main reason for this lack of evidence is, as it should be apparent from the total number of examples, that we do not have a sufficient number of conflicts between any two rules for a hierarchy to emerge; at least not for the corpus we have so far been working with.

We do feel, however, that we have taken a step in the direction of widening the prospects of universality in the phonological area.

*

Having looked at phonological constraints, let us now turn to the s e m a n t i c f a c t o r s in freezing. It is in this area that, to the great regret of many researchers, the picture so promisingly bright for universality in the phonological area, seems to turn bleak. Not that the apparent lack of semantic

[16] "Complex constructions, such as proverbs, are more likely to be retained in the general usage of a language if they are constructed such that place 1 elements are grouped together with other place 1 elements . . . and such that the place 1 part of the constructions precedes the place 2 part." Cooper and Ross, 1975, 68.

[17] Ibid., 100. Emphasis mine.

universals in freezing is considered as something very discouraging by all researchers: Szpyra, for instance, writing of "Freezes which are fixed in a different order in different languages" believes that "Such freezes are particularly interesting and significant since they may point to the existence of different semantic principles in different languages implying that the similarities in the first group are merely accidental . . ."[18]

Whatever the basic approach and the actual semantic constraints, it should first of all be pointed out that as soon as we leave the area of reduplication or "near-reduplication" (both suitable for phonological investigations) and turn to idioms and other coordinate freezes, it becomes obvious that on the whole semantic factors overrule phonology. As numbers of sizable and diversified groups of examples show, "if there is anything that can go wrong (from the standpoint of phonology), it does". In other words, whenever it appears necessary for a semantic factor to override a phonological rule or even a group of phonological rules acting (or rather: trying to act) in unison, it mercilessly does so; and this seems to equally apply to both English and Hungarian.[19]

A highly unusual o > ö (near [e]) sequence becomes rigid for example in Hungarian

megs z o k i k vagy *megs z ö k i k* (approximatively: 'make or break'), and a similar force may have been actively operating in

f o g g a l-k ö r ö m m e l 'tooth and nail', although this one obviously has the syllable rule going for it, too.

Similarly in English, there exist relatively rigid pairs that go against basic phonological predictions, as for instance

husband and wife, or
brother and sister,

both victims of a semantic rule that says Male first.

[18] Szpyra 1983, 36.

[19] It seems that there are a few exceptions. One is English *dead or alive*, which goes against at least one, and possibly two, semantic constraints: 1. Living or animate first, and 2. Positive first; and which may thus be constrained by the powerful syllable law. In Hungarian, which has *élve vagy halva* ('alive or dead', *élet-halál* and *életre-halálra* ('for life and death'), phonology and semantics complement and seemingly reinforce each other — *élve vagy halva* is, according to some bilingual Hungarian American speakers, more rigid than *dead or alive*. Similarly, *back and forth* violates the rule *Front first*, and apparently follows our newly identified low-back pattern; the same phonological rule may be at least partly responsible for strictly ordered Hungarian *éjjel-nappal* ('night and day'), as compared to very loosely ordered English *day and night/night and day*. But even here, the principle *Positive over negative* rears its head where it is perhaps no longer expected; consider Hungarian (strictly ordered) *Nappalok és éjszakák*, confirming to either Russian *день и ночь* or German *Tag und Nacht*, both "day and night" and both ordered.

Natural chronology seems to be another forceful semantic constraint overriding everything else in most cases (with *Now* being perhaps the only occasional rival to it; see below) as e.g. in Hungarian

volt, nincs,

approximately 'it's gone' or 'it's all gone' (lit.: 'there was, there is not'), while

van vagy nincs? — 'Have you got one (any, etc.), or not?'

satisfies the rule Positive first.

This remarkable strength of semantic constraints would, and should, of course, not by itself jeopardize the search for semantic universals in freezing, which is really the most controversial issue. In fact, one could almost logically predict a state of affairs pointing in just the other direction. That this is not entirely the case, however, becomes especially clear when we come to the type of striking examples as e.g. the following:

Hungarian *csont és bőr* — 'skin and bones' (lit.: 'bone and skin');

or English *hammer and sickle* — despite Hungarian, and, what is perhaps more interesting, Russian, 'sickle and hammer':

Hungarian *sarló és kapalács*, and
Russian *серп и молот* } ("sickle and hammer")

In the case of the Hungarian example, phonology does not seem to have had any decisive freezing effect, one way or the other, yet we have a very rigid order as it goes not only against a normally predictable *cross-linguistic analogy*, but also at the same time against two phonological rules, each one by itself a powerful constraint. The only plausible explanation that comes to mind in this case is a type of *interlinguistic analogy*, as predictable for instance from:

English *hammer and tongs*, and
hammer and nails.

But then again, where do those two come from?

Another case of unpredictability is presented by
Hungarian *hideg-meleg* and ("cold and hot")
English *hot and cold*

where we do not have any strong phonological constraints in either case,[20] yet both pairs are strictly ordered — in precisely the reverse direction.[21]

[20] In *hideg-meleg* (*hideg vagy meleg*), we do have an i>e sequence in the first syllables of place 1 and place 2 elements, respectively, but according to all our observations, a single — and not exclusively operating — phonological constraint like this does not normally override a strong semantic rule such as Positive first.

[21] Some colleagues, mainly through verbal communication, have tried to account for some of these opposites in terms of "priorities of values inherent in the nature (or structure) of a given society", or, to put it more simply, in cultural terms. While we do not know how useful this type of explaining may turn out to be, it seems to us that it would go beyond the relative coherency of this particular study.

Considering the above examples, it is less surprising then that, despite a general superiority of semantics over phonology, there are a number of instances of reverse ordering concerning English and Hungarian, supported — possibly — by phonology in one or the other of our two languages. Such examples include, among others, the following:[22]

English	Hungarian
supply and demand	*kereslet és kínálat* ("demand and supply")
nouns and verbs	*ige, főnév* ("verb, noun")
prince and pauper	*koldus és királyfi* ("pauper and prince")
laugh and cry	(*nem tudtam, hogy*) *sírjak vagy nevessek* ('I didn't know whether to cry or laugh')
past, present and future	*jelen, múlt, jövő* ("present, past, future").
(*it was raining like*) *cats and dogs*	*kutya-macska* (*barátság*) ("*dog-cat*" *friendship* -'live a cat-and-dog life')

	English	Hungarian
	summer and fall	*télen-nyáron*
	summer and winter	—('in winter and in summer') — all the year round
	black and white	*fehéren-feketén* ("in white and in black" — in naked reality; there is no hiding it)
but:	*black-and-white* (*tv*)	'fekete-fehér' (tv)
	buy and sell	*adásvétel*, ('selling and buying') *adás-vevés, adok-veszek* ('peddling, wheeling and dealing, huckstering')

However, our relatively lengthy list notwithstanding, we cannot suggest systematic patterns of differences between English and Hungarian at this point,

[22] We do not suggest that there is a particularly prominent contrast just between English and Hungarian. Similar lists could be constructed for English versus other languages, but so far they do not seem to have generated any significant explanatory power, although they make the picture perhaps more interesting, which is, it must be admitted, the single major reason for their inclusion here.

As one direct continuation of the present study, *a group of languages*, including related languages as well as non-related ones, will be examined. This study in preparation is expected to yield some more insight into some of the details of semantic ordering and also concerning the whole complexity of the freezing phenomenon.

not even in spite of the fact that individually, the majority of the cases could more or less be accounted for on either semantic or on phonological grounds, or as a combination of the two.

In addition, and despite all the differences, the picture seems better than absolutely hopeless for similarities. This is obvious when we look at the dozen and a half or so s e m a n t i c d o m a i n s Cooper and Ross identified, that is for which they have found freezes in English — an average of nearly a dozen for each domain. In the vast majority of the cases where there is a comparable Hungarian freeze (a word-for-word equivalent, a near-equivalent or if only a similar construction), t h e o r d e r i n g i s t h e s a m e i n H u n g a r i a n, not infrequently at the price of violating established phonological rules. For a brief illustration consider the following select list — (the whole collection could — although should of course not — be mistaken for a group of "mirror translations" or "loan translations"):

	English	Hungarian
H e r e:	*here and there*	*itt vagy ott, itt-ott*
	come and go	*jön-megy, jövés-menés*
	this, that and the other	*ez, az, amaz* etc.
M a l e:	*man and woman*	*férfi és nő*
	husband and wife	*férj és feleség*
but:	*ladies and gentlemen*	*hölgyeim és uraim* ('ladies and gentlemen')
	bride and (bride)groom	*menyasszony és vőlegény* ('bride and groom') —

— freezes representing politeness conventions in both languages.

P o s i t i v e:	*positive or negative*	*pozitív vagy negatív*
	plus or minus	*plusz vagy mínusz*
	more or less	*többé-kevésbé* etc.
N o w:	*sooner or later*	*előbb vagy utóbb*
	now or never	*most vagy soha*
	yesterday and the day before (yesterday)	*tegnap, tegnapelőtt* etc.
L i v i n g:	*life and death*	*élet és halál*
	live or die	*élni vagy (meg)halni* etc.
G e n e r a l:	*form and substance*	*forma és tartalom*
	general and particular	*általános és különös*
	word and deed	*szó(val) és tett(el)* etc.
P o w e r s o u r c e:	*horse and rider*	*ló és lovas*
	gin and tonic	*gin tonikkal* ('gin with tonic')

In naming mixed drinks, neither English nor Hungarian seem to show any exceptions from the apparent rule that the alcoholic ingredient must take first place; further examples include

Scotch and soda	*whisky szódával/kólával*
bourbon and coke	(*'whisky with soda/coke'*)

The Food and Drink Hierarchy (approximately: fish > meat > drink > fruit > vegetables > baked goods > dairy products > spices) is the largest domain Cooper and Ross identify and for which they list more than 30 examples. Yet, for about two thirds of these freezes there are no one-to-one Hungarian correspondences, one reason for which is indeed definitely cultural (unlike perhaps the cases of *hammer and sickle, prince and pauper* etc.), given the fact that many of the diches, spices etc. themselves are different in the two culture areas.

Another reason for the lack of Hungarian equivalents in some cases is grammatical: suffixational relations for instance ab ovo determine the nature of many binary constructions which therefore come about by means other than freezing. Thus we have e.g. Hungarian *vajas kenyér, zsíros kenyér, sajtos kenyér* etc. — 'bread and butter' (lit.: 'buttered bread'), 'bread and drippings' (lit.: 'greased bread'), and 'bread and cheese' (lit.: 'cheesed bread') etc., respectively.[23]

But even in this domain there are some "mirrors" and "near-mirrors", including

English	Hungarian	
fish and game	*hal és vad*	(lit.: 'fish and game')
meat and potatoes	*hús* krumplival	(lit.: 'meat with potato')
ham and eggs	*sonka* tojással	(lit.: 'ham with egg')

and note even:

bread and water	*kenyér és víz*

a perfect "mirror" with both expressions going against the established hierarchy.

And what is perhaps more significant, of the thirty or so English freezes, in not a single case have we found a Hungarian freeze with the reserve ordering.

[23] This *attribute+noun* type construction is muchm ore frequent in Hungarian than in English. Similar sets of expressions are constructed with a number of other nouns, notably those denoting types of food, as *tészta* — 'noodles', 'pastry', 'vernicelli'; *leves* — 'soup'; *kalács* — (a kind of white bread); *főzelék* — (vegetables made with shortening); and so forth. The numerous attributes combined with these nouns prevent the formation of many *bread and butter* — type binomials in Hungarian.

To add one very convincing example to our long list of similarities: the ordering of the cardinal geographical referents and all their sub-components (North, South, East, West, Northwest, North by Northwest etc.) is strictly and rigidly the same in Hungarian as in English, even at the price of having to violate the otherwise powerful syllable rule in:

Észak-dél lit.: 'North-South'

On the other hand, even these (and many other similar) examples are altogether little solace for the universalist, who should of course be also aware, for instance, of German and Spanish that have *West(en) und Ost(en)* and *del sur al norte*, respectively — note the possibility of a phonological explanation in both cases.

At one point, Hungarian seems to conspire with Yiddish to ruin an otherwise very uniform-looking cross-linguistic picture in space-axis ordering,[24] and what is worse, there is no apparent explanation — phonological or whatever — for the conspirancy; cp.:

Yiddish	*orop un aroyf*	("down and up")
	unten un oybn	("below and above"); and
Hungarian	*le-fel* (*föl*)	
	lefelé vagy fel/föl/felé	(both "down and up" — adverbs of direction),
	lent vagy fent	("down or up" — adverb of place).[25]

For another group of related freezes as examples of ordering in the reverse direction compared to English, consider

	ki-be	("out and in" — adverbs of direction);
	kifelé vagy befelé	
	kint vagy bent	("outside or inside")
	kívül-belül	('outside-inside' — 'inside out');
and also	*körülbelül*	('around and within' — 'approximately'

[24] Consider, among others, the following:

Latin	*sursumac deorsum*	('high and low')
German	*auf und nieder*	(both 'up and down')
	auf und ab	
	oben und unten	('above and below')
	über und unter	('over and under')
Indonesian	*naik turun*	('ascend and descend')

— examples given by Cooper and Ross 1975, 87.

[25] It must be admitted that the Hungarian contribution to the conspiracy is somewhat half-hearted: none of the three freezes are very rigid, in fact, there are native speakers who say that they are hardly ordered at all; that is, *föl-le* and *fönt-lent* are just (or almost) as often used. In addition, there is the idiom *fel s alá* (*járkál*) '(walk) up and down'), while on the other hand this may be (partly) counterbalanced by loosely ordered *l e is út, f e l is út* (expression of sending someone away angrily).

Apart from Yiddish and Hungarian, we are aware of Russian *взад и вперёд* ('down and up'), and are almost certain that further "conspirators" could be found.

which may again be rare examples of a phonological rule (or possibly two: there is also "labialization" in place 2 elements) operating successfully against semantics, or at least against the ordering in English.

And the worst is perhaps yet to come. At the end of their paper, Cooper and Ross suggest two potential universals that they " . . . have not yet been able to shoot down"[26] — and which are as follows:

a) Star-Extra: Mick Jagger and the Rolling Stones,
Van Cliburn and the Moscow Philharmonic,
John Wayne and a cast of thousands,
Snow White and the Seven Dwarfs; and
b) Chronology: in a freeze of two verbs which are intended to be in a temporal sequence the place 1 verb denotes the earlier action."[27]

Universal "a" is confirmed by Hungarian. For all four freezes quoted, Hungarian has its perfect formal equivalent (or translation), and so it does for a large number of similar pairs in English without any exceptions that we are aware of. In addition, it has of course its own numerous examples of "Star-Extra", including, among many others:

Lakatos Sándor és népi zenekara (S á n d o r L a k a t o s[28] and his Gipsy Band') and

Zalatnay Sarolta és a három Tini (a pop group).

Universal "b", however, is in trouble: for what seems to be a painful but unavoidable instance of universal-shooting, consider

Hungarian *jár-kel*

(*jártában-keltében* etc.), an absolutely irreversible pair meaning 'wander around', 'be on the move', 'come and go', 'walk about', 'travel about' etc., where both *jár* and *kel* are otherwise separate, individual verbs, *jár* meaning 'walk' and *kel* meaning 'get up' or 'rise'.

[26] Cooper and Ross 1975, 102.

[27] Ibid: 102.

[28] Given the subject matter of this paper, it is perhaps not entirely inappropriate to note at this point that Hungarian, unlike most other European languages, orders personal names in the reverse direction, i.e. first names last, last names first. This goes traditionally and without exception for H u n g a r i a n personal names, while the order for non-Hungarian ones changed historically with the changes in approaches and attitudes toward translation. According to prevailing translation practices in the 19th century and earlier, e v e r y t h i n g was to be translated (or often rather: adapted); by now this approach has turned into what we might call a minimalist one. Thus, for instance, the full "Hungarian name" of Shakespeare has changed from earlier L á n d z s a r á z ó V i l m o s (lit.: 'Spear-shaking William') to a somewhat more plausible *William Shakespeare*. However, many historical names have kept the characteristics of an original translation including the reversed order. Thus we have *Oroszlánszívű Richárd* ('Richard the Lionhearted'), *Rettegett Iván* ('Terrible Ivan') and so forth.

By way of a tentative conclusion it must be noted that in this study we have barely gone beyond scratching the surface of a large and complex field called the freezing phenomenon. Some of the aspects of research not discussed here but which should eventually be incorporated in further studies include r h y t h m,[29] v o w e l h a r m o n y for Hungarian (and other agglutinative languages) — vowel qualities change very rapidly in respect to consonants which may lead to an explanation of a number of "exceptions", and shed more light on the interaction of phonology and semantics; and also a general h i s t o r i c a l l o o k b a c k. Also, as we mentioned it earlier, the investigation should be extended over units larger and more complex than simple binary pairs, as there are indications that the freezing phenomenon goes beyond binomials. Finally, it seems necessary that we include other languages in future research, thus making a comparison of related languages as well as non-related ones, and so hoping to gain more insight in general and more knowledge of H u n g a r i a n freezes in particular.

References

Abraham, R. M.: Fixed order of coordinates: A study in lexicography. In: Modern Language Journal 34 (1950), 276—287.
Cooper, W. E. and Ross, J. R.: Word order. In: Grossmann, R. E., Stan, L. J. and Vance, T. J. (eds): Papers from the Parasession on Functionalism. 1975, 63—111.
Gonda, J.: The functions of word duplication in Indonesian languages. In: Lingua (1949—50), 2.
Gustaffson, M.: Binomial Expressions in Present-Day English. Turku 1975.
Jespersen, O.: Growth and Structure of the English Language. Leipzig 1905.
Luelsdorff, Ph. A.: Some modal freezes. In: Anglistik und Englischunterricht. 8: Semantik. Trier 1979.
Malkiel, Y.: Studies in irreversible binomials. In: Lingua 8 (1959), 113—160.
Molnár J.: A magyar beszédhangok atlasza. Tankönyvkiadó, Budapest 1972.
Oden, G. C. and Lopes, L. L.: Reference for order in freezes. In: Linguistic Inquiry 12 (1981), 673—679.
Pinker, S. and Birdsong, D.: Speakers' sensitivity to rules of frozen word order. In: Journal of Verbal Learning and Verbal Behavior 18 (1979), 497—508.
Scott, F. N.: The Order of Words in Certain Rhythm-Groups. In: Modern Language Notes XXVIII. (1913), 237—239.
Szpyra, J.: Semantic and phonological constraints on conjunct ordering in English and Polish. In: Kwartalnik Neofilologiczny XXX (1983), 34—53.
Thun, N.: Reduplicative Words in English. Uppsala 1963.
Vapordschiev, V.: Parallelen und Unterschiede in den Zwillingsformen im Deutschen, Russischen und Bulgarischen. In: Wiss. Zeitschrift der Humboldt-Univ. zu Berlin XXIV (1975), 387—390.
Vapordschiev, V.: Wortpaare im Bulgarischen und Deutschen ohne syntaktisches Bindungsglied. In: Linguistische Arbeitsberichte 20/21 (1978), 285—293.

[29] We are not the first to suggest the inclusion of the concept of rhythm in the study of binomials. As early as about eighty years ago, Jespersen (1905) initiated a rhythm theory. However, he was later criticized e.g. by Scott (1913) and, to the best of our knowledge, the study of irreversibility and rhythm was subsequently largely abandoned.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 181–195 (1986)

THE SEMANTIC ORGANIZATION OF WORD-FORMATION PARADIGMS AND DIACHRONY

AMANDA POUNDER

(University of Vienna, Austria)

1. Introduction

1.1. There can be observed certain serious contradictions in the discussion of semantic relations between words related through word-formation processes. For example, it has been stated in various frameworks[1] that the formation, always by rule, of a word is automatically prevented by the presence of a synonym having the same base. How is such a statement to be reconciled with typical observations in the context of historical descriptions of individual languages to the effect that word-formation synonymy can be created both by rule and by lexical processes, although unstable and tending to decay? Besides questions of relative chronology, we have here a theoretical contradiction as to the existence of paradigmatic synonymy in general and to the manner of its creation as well. In order to evaluate the above and related claims, a differentiated conception of the semantics of word-formation is necessary. In the following sketch of some diachronic aspects of the paradigmatic relations of polysemy and synonymy, I shall try to satisfy this minimum requirement.

1.2. Concepts and definitions. The following concepts are essential to the framework on which the present discussion is based.[2]

1.2.1. Status of word-formation. Inflectional and word-formation morphology together form a separate component of the grammar, that is, are neither part of syntax nor of the lexicon,[3] although of course they interact with both.

1.2.2. Word-formation. In this framework there are two series of word-formation rules, namely semantic rules and formal rules, which are paired

[1] E.g. M. Aronoff 1976, p. 55; F. Plank: Morphologische (Ir-)Regularitäten. Tübingen 1981, 148f. H. Paul: Prinzipien der Sprachgeschichte, Halle 1909.

[2] These and other related concepts are presented in detail in my dissertation. (Pounder, 1987).

[3] The lexicon is considered to be the (static) store of all of the speaker's semantic, syntactic, referential, and extra-linguistic knowledge of lexical items (words, stems) and should be modelled in as psychologically realistic a manner as possible; this implies a high degree of flexibility and plasticity, as well as a high degree of redundancy.

with each other (= word-formation operation) in various combinations to form or to analyze words.

1.2.3. Word-formation paradigms. The word-formation paradigm is a construct showing formal, semantic, and other relations between (simplex and complex) stems and words (stems) built from them, and as such performs a bridging function between the morphological component and the lexicon. It is meant to correspond to psychological reality, that is, it is certain that one of the modes of organization that a speaker makes use of is the relaying of elements related through word-formation. On the other hand, the graphic representation chosen has of course no "real" counterpart, so that this may be determined by the requirements of the problem at hand (e.g. the degree of abstraction of the representations of morphological units or of processes).

The word-formation paradigm is both static (where it might be equated with the notion of word-family or word-nest) and dynamic (processual) in nature.[4] This means that every (ad hoc or other) word-formation act is carried out in relation to a word-formation paradigm[5] (cf. J.v. Marle's reference to *elements in absentia*), including the spontaneous binding of other formal and semantic rules than in the "lexically fixed" pair.

There is a high degree of individual variation in the formal and semantic organization and content of word-formation paradigms; in this respect, they might be compared to the lexicon.[6]

1.2.4. Semantics of word-formation. 1.2.4.1. Word-formation rules. Although arbitrarity can probably not be completely avoided, our goal should be to propose word-formation rules which can be empirically proved to be "realistic" and that are not so specific as to be controversial.[7] These should express standard (not language-specific) relations, whose application is dependent on the semantic content of the base on the one hand and whose identification is aided by the context on the other hand. Non-relational elements of meaning are considered to be lexical; for example, for a complex

[4] L. Guilbert (La Créativité Lexicale, Larousse, Paris 1975) was probably the first to develop the idea of the dynamic word-formation paradigm. Cf. also J. v. Marle 1984.

[5] It could be that certain kinds of word-formation are not necessarily directly paradigmatic in nature, e.g. ad-hoc compounding or syntactic derivation in text where the immediate, cohesion-forming relationships are between preceding or following lexical elements.

[6] Cf. S. Henderson-Taylor (On the Acquisition and Completion of Lexical Entries In: Papers from the Parasession on the Lexicon. Chicago Linguistic Society 1978, p. 347–356)

[7] Cf. here Aronoff 1976, 1984; W. Zwanenburg 1984, R. Beard 1981, B. Warren 1978, M. Ljung 1970, I. S. Uluxanov 1977. Pounder 1987 contains a fuller discussion of word-formation semantics as well as a complete system of semantic rules.

word designating an object, the closer specification of the nature of the object ("human", "non-human", "habitual", "non-habitual" etc.) would not be part of the rule, but rather be lexically or contextually determined. The idea that word-formation can regularly express elements of meaning that are not part of the word-formation meaning appears at first paradoxical, and indeed conflicts with our superficial intuitions as speakers of a language; however, it must be borne in mind that these rules state relations between words (or their referents) and do not describe things.[8]

A simplified example of a possible semantic rule appears below:

'Y' = 'LIKE ('X')' for the comparison relation,
e.g. 'doughy'$_1$ = 'LIKE ('dough')'
'milk-' = 'LIKE ('milk')
in *milk-white*

where 'Y' is the word-formation meaning of a complex word (stem), LIKE the label of the function or relation, and 'X' represents the stem content.

1.2.4.2. Synonymy and polysemy. There has been a good deal of discussion concerning word-formation synonymy, e.g. criteria such as substitutability, the question of whether absolute synonymy is possible, and so on. This is among other things due to the fact that often no effort is made to analyze the semantic structure of complex words, so that elements of lexical meaning get mixed up with the general abstract rules just mentioned. I will not discuss the matter further here, but will define w o r d-f o r m a t i o n s y n o n y m y as follows:

Two words are (fully) word-formationally synonymous when:

1) the formal requirements are satisfied (i.e. same or related base, same lexical class (exceptions here) and so on), and

2) they share the same semantic rule or rules.

Likewise, partial synonymy occurs when the words have at least one semantic rule in common and at least one not; inclusion occurs when one word has all semantic rules of the other and at least one additional one, etc.

Thus it is possible for two words to be word-formationally synonymous but n o t lexically synonymous, i.e. not substitutable in all or even any contexts.

W o r d-f o r m a t i o n p o l y s e m y: a word is word-formationally polysemous when its form corresponds to more than one semantic rule. It must then be lexically polysemous, of course.

[8] There are possible exceptions to this, such as the treatment of "feminizing" suffixes. In general, it can be said that word-formation rules differ as to degree of abstractness, that is to say, are hierarchically ordered (cf. W. Zwanenburg 1984, I. S. Uluxanov 1977).

We therefore recognize a semantic structure of complex words, which allows us to show several levels of synonymy and polysemy.

1.2.4.3. Levels of meaning. A complex word may be said to have four levels of meaning:[9]

1) word-formation meaning: the semantic rule or rules corresponding to the formal rule.

2) lexical meaning: the information available to the speaker that is not contained in the word-formation rule; it refers to the whole word or its referent and is not necessarily relational. The treatment of lexicalization on a level separate from word-formation prevents the common confusion with lexicalized formations: it is often forgotten that the semantic rule is extractable (and is extracted) from the total meaning even in the presence of lexicalization.

3) situational role: This notion is adapted from I. A. Mel'čuk's syntactic roles; in some ways it is comparable to the thematic roles of other frameworks. The situational roles correspond to the participants (actants) in a given pragmatic situation (in our case, that referred to by the content of the stem). This level is relevant for all lexical classes, but more particularly for nouns and verbs.[10]

4) stylistic level: This is the most superficial semantic aspects concern especially function as intensifier and secondary semantic rules.

Thus can the meaning of every complex word be specified according to these four levels. It may be, in a given case or set of stems/paradigms, that the one or the other of these levels does not happen to be relevant.

On the basis of these concepts then I shall examine the problems mentioned in **1.1.**, namely paradigmatic polysemy and synonymy from a diachronic point of view.

2. Polysemy in the word-formation paradigm

2.1. Word-formation polysemy. This semantic relation concerns the relationship between the base of a word-formation and the word-formation

[9] The number four was arrived at empirically. It may be verified by observation of diachronic development: the levels described here are those on which semantic change occurs. Also, when comparing two or more members of a paradigm, we see that these are precisely the aspects in which they differ from one another.

[10] It is conceivable that this level should not constitute a special level, but rather be included in the lexical level. The reasoning for this would be that as the lexical level is "responsible" for the referential distribution of complex words, the fact that a given word refers to one or the other participant in a given situation is a lexical question. Others would be inclined to characterize the situational roles, as their originator did, as being syntactic in nature. While admitting these possibilities (and favouring the former), I prefer to consider the situational role level as separate, though of course less general than the lexical level.

itself, that is to say: the semantic content of the stem is used as the basis for more than one semantic relation (rule). For example, from a noun denoting a human being may be derived the relations of possession, of origin, of comparison, of the related "suitable to", as well as the most general rule expressing relatedness. The development of word-formation polysemy occurs when the number of semantic rules expressed by a given form is less than that of the possible semantic rules regularly expressed by word-formation processes in a given language for each subset of the lexicon. In fact, even apart from the "filling up" of the potential spaces for semantic rules in a particular word formation paradigm, word-formation polysemy appears to arise very naturally. Many semantic rules are conceptually related to others, as has often been remarked, so that the possibility of word-formation polysemy expanding is high.[11]

An example of semantic expansion on this level is the German paradigm

KNOT-:	KNOT-	*knotig*	*knotig*
		LIKE	LIKE
			WITH
		17th—18th c.	Mod. G.

Early dictionaries only give (the equivalent of) the comparison relation; in the most recent Duden the "having" relation is given as well ('WITH').[12]

A natural question is then: what can check the development of word-formation polysemy in one case or another? If one subscribes to the principle that the ideal Form-Meaning relation is biunique, then one should expect word-formation polysemy to be inhibited, or, should it be allowed to arise, to be unstable, that is, to be eliminated after a short period of time. There really does not seem to be any such principled check on the coming into existence of word-formation polysemy, if this polysemy is "normal" in the word-formation system or subsystem in question. Furthermore, a "division of labour"between different forms, should it occur, is itself unstable.

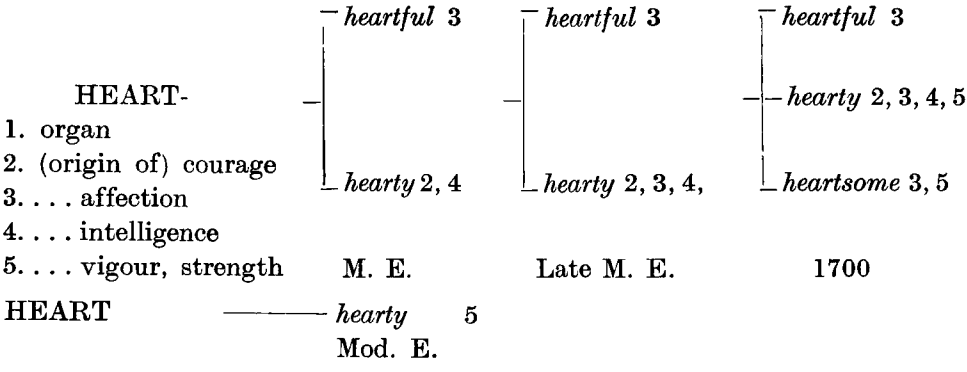
[11] The potentiality of word-formation polysemy can be easily tested. For example, I tested reactions to regularly built forms in usual textual contexts going back to "unusual" semantic rules ("unusual" with respect to the individual paradigm) for German. There was, as was to be expected, great individual variation, but the results very clearly showed a considerable tolerance for these different "potential" rules. Another proof of this inherent polysemy was a test in which the subjects were asked to build words from stems for which there is lexically speaking no derivational paradigm. The variety of semantic rules used within the test group was surprisingly great.

[12] This is really a lexicographical question; due to the potentiality of word-formation polysemy it seems natural to suppose that *knotig* in the sense of e.g. "knotty wood" ('WITH') was possible earlier.

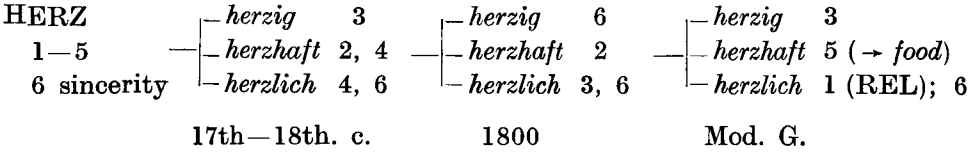
This and the following are very concrete representations of paradigms, to avoid complicating matters through unfamiliar formal symbolism.

2.2. Lexical polysemy. Polysemy on this level can take several forms, the most important of which is the variety of referents "covered" by a given formation (cf. English *sleeper*: person, object (train car, child's garment, among others; cf. Dressler 1980); in cases where lexicalization has taken place, it is possible for a formation to have a "straight" (i.e. purely word-formation) meaning a n d the more specific one at the same time (this sort of polysemy is always potential in text); in connection with this arises the question of how much of the referential content is met with in the base (for example, the formation may refer to only one aspect of this content; denominal adjectives with the semantic rule of comparison may refer to only one quality or characteristic of the basis of comparison, e.g. with substances to taste, colour, consistency, or to more — *nudelig* ('noodle -y') could in principle be many things, but refers only to consistency (at least according to a small number of informants), whereas *kreidig* or the Engl. equivalent *chalky* can refer to colour, taste, and so on.)

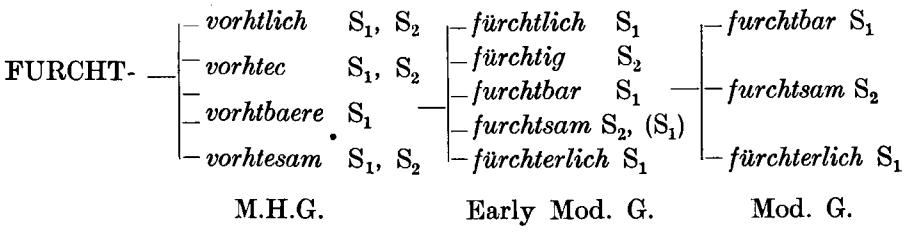
Now, change is possible — and frequent — with all aspects of lexical polysemy. The following examples show reduction and expansion in a fairly complicated paradigm, with lexicalization.



With the many lexical meanings of *hearty* and earlier *heartful* — 'intelligent', 'strong', 'courageous', 'affectionate' etc. — and the different referents — originally used for persons only, the words were extended to such things as food and drink etc. — there has been only the one semantic rule, namely 'Y'='WITH('X')' ("having" -relation). In the German example below, there are more semantic rules available, namely 'FROM' (origin relation', 'REL-' relatedness relation) etc.

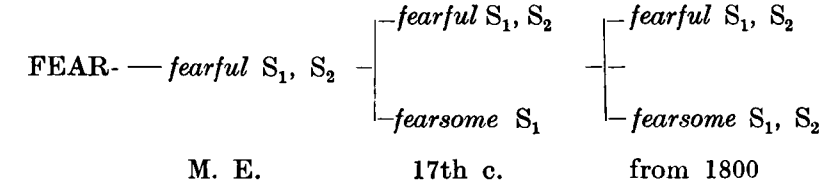


2.3. Polysemy of situational role. Change in the polysemic relations of a word-formation to its base can often be observed at this level. With both verbs and adjectives we see some regular lexically dependent splits, which appear to be often difficult to uphold. The following example shows how two different languages "deal with" equivalent bases:
German: (reduction)



where S_1 = source (a sorrowful event)
S_2 = experiencer (sorrowful person)

Here we see a reduction in the number of possible situational roles from Middle High German onward; a "division of labour" is effected. This would appear to confirm the notion of biuniqueness being optimal, except that there many examples in German of paradigms where polysemy is present. The English example shows expansion within the paradigm:



2.4. Stylistic polysemy. Here we see expansion and reduction of plural stylistic values (e.g. intensifier function and "literal" meaning (neutral) at the same time), e.g. *fürchterlich* ('causing fear') acquired soon after its appearance the intensifier function, probably in analogy to *furchtbar*, already polyvalent.

2.5. Stability of paradigmatic polysemy. The question arises of whether it is possible to show that one level is more susceptible (or resistant) to change than another. It would in fact appear that the meaning levels are not equivalent in this respect: although the existence and growth of polysemy on all levels are very natural and frequently occurring phenomena, they are most predictable on the word-formation level. Thus it is difficult to find examples, even with a very large corpus, of reduction and

expansion within the same conditions of regular Form-Meaning rule combinations on the first level, and what there is seems to be of the nature of a lexicographical accident. The predictability of word-formation polysemy is due to the fact that the set of possibilities is finite; for a given word-stem, the potential polysemy is a function of pragmatic possibility dependent on stem content, and the regular Form-Meaning rule combinations for the relevant subset of the lexicon. As well, we must not forget that these word-formation rules are such that they are known, that is to say, psychologically real to the speakers/hearers, and are so to speak automatically extracted and applied, just as formal elements are, so that a speaker/hearer can very well judge which semantic rule is meant contextually, just as he can interpret a compound word. That is, there is i n p r i n c i p l e no need for monosemy here, and normative efforts in this direction may very well be rejected.

As far as limits on polysemy are concerned, it is difficult to generalize. as word-formation polysemy is not uniformly tolerable over a given system (cf. in German the restricted number of semantic rules for animal names as compared to bases denoting human beings); in addition, one must not forget that in some cases only one semantic rule may be pragmatically imaginable.

The discussion of paradigmatic polysemy can be concluded with the suggestion that statements as to the undesirability of polysemy as an asymmetrical relation must be relativized. There seems to be no need to express every semantic rule or every nuance of meaning with a separate lexical unit; historical developments show a preference for polysemy over large paradigms, that is, ones with many members of the same order. The reasons for this may be that expansion and reduction on the word-formation level have a different status: here we have formation by rule, whereas expansion and reduction on the lower lexical levels are due to lexical processes.

3. Synonymy in the word-formation paradigm

As indicated in **1.1.**, paradigmatic synonymy is treated in an undifferentiated manner, generally speaking. As a corollary to the statement that lexical synonymy does not exist, paradigmatic synonymy is often declared to be impossible, whereby there is usually no effort made to consider any other than the lexical meaning of complex words. Thus it was possible to derive a principle ("blocking principle" etc.) saying that the presence of a complex word with a meaning X prevented the formation of a new complex word with that same meaning, this principle being used to explain limits on productivity.[13] The main problem here, apart from the vagueness of the semantics, is that

[13] Cf. e.g. M. Aronoff 1976.

the principle runs contrary to fact. Others,[14] confronted with the existence of paradigmatic synonomy in actual cases, adapted this principle, saying that a truly ("syntactic") productive rule was not subject to it, so that synonomy could come about; this is quite a crass contradiction.

We must consider the "blocking" of a word-formation act as functioning momentarily only; it cannot principally exclude the formation of a synonym or of the application of a new semantic rule to an existing form taking place in five minutes, next week, or next year. "Blocking" of synonymy involves the activation of the normative capacity of the speaker/writer and can thus only apply to individual cases. Cf. here J. v. Marle's expression "non-systematic non-occurrence" of rule application and his statement that (p. 205): "blocking is a phenomenon which we consider to relate above all things to 'language use' (and not to language structure)".

There arises the question of whether the "blocking" or the reduction of paradigmatic synonymy applied to the word-formation level and to the lexical levels equally. If it could be shown that this is not the case, that the special status of the word-formation level observed with paradigmatic polysemy has a parallel with the semantic relation, then one could speak of a certain potentiality existing here as well.

3.1. Word-formation synonymy. The first possibility for the development of word-formation synonymy is the application of a rule or rather rule combination when a paradigm member of the same order and with the same semantic rule(s) is present. It may be that the first form is no longer as frequently paired with the(se) semantic rule(s) as formerly, so that the new form-meaning-combination represents the newer, more typical, normal relation. The creation of word-formation synonymy by rule can also occur when there is no such difference between the normality of the rule pair. The possibilities of combination correspond to the scope of each rule with respect to the meaning of the base (class of bases). For example, in German, for denominal adjectives whose bases denote concrete objects or substances, the semantic rule expressing comparison may be paired with the suffixes *-ig* und *-haft*, whereas the rules combining *-lich*, *-isch*, *-haft*, and *-ig* with the stem are paired with the same semantic rule for stems denoting human beings or abstract notions. In addition to this, the same semantic rule can be paired with little restriction with formal rules of several word-formation processes, e.g. derivation, compounding, serial compounding, participle formation, etc. Examples of such paradigm growth could be seen in the above FURCHT-, FEAR-, HERZ- and HEART- paradigms; below are two more German paradigms:

[14] Cf. e.g. E. Broselow; also the discussion in R. Schupbach.

a) BIEN- 'bee' (test example)
- *bienenhaft* (dictionary)
- *bienig* 'bee-y') (test answer)

Semantic Rule: 'Y' = 'LIKE ('X')'

b) KLATSCH- 'gossip'
- *klatschaft* previously existing formation
- *klatschig* 'gossipy' new formation (Duden 1976)

Semantic Rule: 'Y' = 'WITH ('X')'

Further, when several forms are already present in the paradigm, and they were on this level, not synonymous up till the point in question, a new semantic rule may be applied (analogical extension). For example, the German METALL- ('metal') -paradigm was split semantically for the rules OUT OF (*metallen*) ('metal-') and LIKE (*metallisch*) ('metallic'); these are now partially synonymous as *metallen* can be used with the semantic rule LIKE (and *metallisch*) is also word-formationally polysemous, in that the form is paired with REL (relatedness relation) as well.

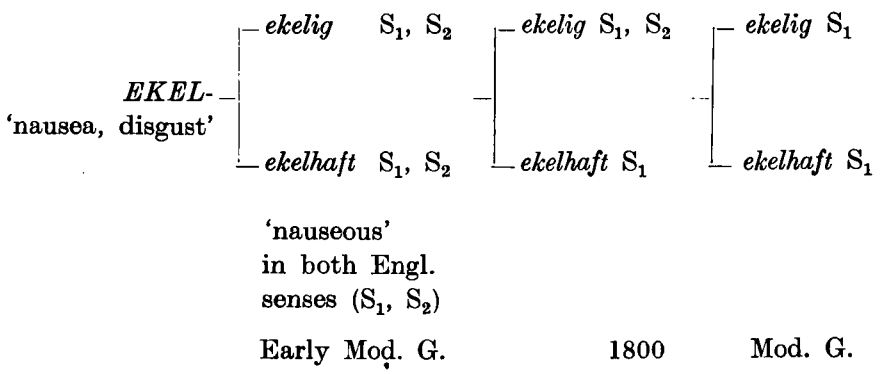
These processes can be subsumed under the general concepts of extension of context or of unchecked rule application; furthermore, they are parallel to those typical of the development of polysemy in the word-formation paradigm.

3.2. Lexical synonymy. Synonymy may arise through neutralization of a former nuance or contextual differentiation, or again through natural analogical extension of the one or the other paradigm member (through the inability to "keep the two forms apart" or remember "which is which"); this is encouraged by contexts where the original differentiation is irrelevant, or new contexts, where both words may apply. An example of neutralization is provided by the German MEISTER- paradigm ('master'): the early Modern German words *meisterhaft* and *meisterlich* were both formed by the semantic rule 'LIKE', but referred to different characteristics of a 'Meister'; in modern German, both have the meaning 'masterly'.

MEISTER- < *meisterhaft* ('lordly, arrogant') / *meisterlich* ('masterly') < *meisterhaft* ('masterly') / *meisterlich* ('masterly')

An example of expansion: the English words *observation* and *observance* were originally derived from different lexical meanings of the verb *observe*; Fowler gives examples of each derivation being used in both senses (p. 395) "to reinforce *observance* with imagination" "The British Govt. has failed to secure the *observation* of law . . ."

3.3. Synonymy of situational role. The extension of application of a given lexical item to another role in a given pragmatic situation is frequent (cf. H. Fielding (Joseph Andrews) "this sick and painful bed") or the German expression *durstige Baustelle* ('thirsty construction site', i.e. one where one is not allowed to drink beer). Should there exist differences between two forms, this difference can be neutralized, either through extension of the one or reduction of the other, e.g. the FEAR- paradigm above, where we saw the extension of *fearsome*, or the FURCHT- paradigm, where the number of roles for *furchtsam* was reduced. An additional example is the EKEL- paradigm in German, where two stages of total synonymy and one of inclusion may be observed (there is still a lexical difference, but that is not relevant here).



3.4. Stylistic level. Continuing the same pattern, neutralization of differences causing reduction to synonymy or extension through application of a secondary rule can be observed here, for example, in the FURCHT- paradigm, the new member *fürchterlich* came to acquire the intensifier function in additional to the literal sense, thus becoming totally synonymous on all levels with *furchtbar*.

3.5. Reduction of synonymy in the word-formation paradigm. This is a much discussed topic: it is claimed that synonymy is eliminated in one way or another as being superfluous as soon after it occurs as possible.[15]

[15] Cf. e.g. I. Ohnheiser, H. J. Grimm, V. V. Vinogradov, I. M. Mal'ceva i.a.

That a formal reduction ("paradigm deflation") and a reduction of synonymy in the word-formation paradigm is a frequent occurrence cannot be denied, as historical studies of e.g. French, German, or Russian show.[16] One may however legitimately ask if these reductions are exclusively a reaction to "uneconomic" synonymy, or if there might have been "simply" a massive formal reduction, tied to a change in priority of the various functions of word-formation in general. A "desire" for paradigms with relatively few members, although *not* a wish for the elimination of a motivated lexicon ! — would then be behind the principal tolerance of polysemy and to both the existence of and limits on synonymy, i.e. analogical extension of the same form leads to partial or total synonymy as well as to polysemy, as there is reluctance to "create" a new form-meaning combination.[17] As far as polysemy is concerned, it is "better" that one form have more meaning (semantic rules) than that there be several monosemic, different forms. Such preferences are language-specific and can change over time within a particular language.

The processes of reduction are familiar: here follow a few brief examples on each of the paradigmatic levels.

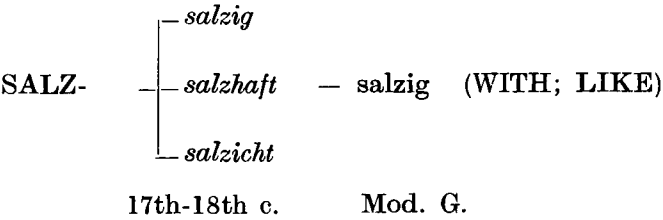
3.5.1. Reduction of word-formation synonymy. As evidence for this interpretation, consider the normative efforts of language critics such as Adelung in the late 18th and early 19th centuries in. Germany: these insisted on the observance of a "functional split" in denominal paradigms between adjectives with the LIKE and the WITH rules, complaining that no one seemed to be able to keep the difference straight; the resulting developments showed the failure of their labours, as the paradigms were systematically reduced to (usually) only one derived member, necessarily polysemous.

Differentiation: Here there may be preferences for certain separations and tolerance for other partial synonymies and polysemies, for example denominal adjectives from bases denoting substances in German: *hölzern* (OUT OF; LIKE) — *holzig* (WITH; LIKE) ('wood'); *gläsern* (OUT OF) — *glasig* (LIKE); *drähtern* (OUT OF) —*drahtig* (LIKE) ('wire'); or another type of separation in Russian, namely between the relative adjective on the one hand and all other semantic rules on the other hand: *vkusovoj* (REL) vs. *vkusnyj* (WITH) (cf. also German *geschmacklich* (new) vs. *schmackhaft, schmackig*; English *taste-* vs. *tasty*).

Formal elimination of one member: Here there are countless examples; cf. the systematic reduction of German paradigms with concrete stems, e.g.

[16] Cf. M. Gawełko, V. V. Vinogradov, I. M. Mal'ceva.

[17] This of course does not constitute an absolute "blocking principle" or mean that word-formation paradigms never become enlarged.



3.5.2. Reduction of lexical synonymy. The originally synonymous *schmerzhaft* and *schmerzlich* (both 'painful', both S_1) are now differentiated according to the type of pain involved: *schmerzhaft* refers to physical pain and *schmerzlich* to emotional pain, so that there is a different expression for "painful wound" and "painful event".[18]

3.5.3. Reduction of synonymy of situational role. Cf. the German FURCHT- paradigm.

3.5.4. Stylistic synonymy. The originally synonymous *kindisch* and *kindlich* ('childish', 'childlike') are now differentiated according to the secondary semantic rule expressing pejorative shading. The situation is further complicated by the third paradigm member *kindhaft* and the capacity of *kindlich* in unambiguous contexts to be matched with the secondary semantic rule.

Synonymy on the word-formation paradigm seems to be more susceptible to reduction than polysemy, perhaps partly due to the fact that the potentiality of fulfilment mentioned above is lacking in synonymy. The tolerance level for synonymy seems to depend on the semantic content of the stem (class) and what is "normal" for this class in a given language. For, example, in German, bases denoting concrete objects seem not to allow many formal rules at once in adjective formation, although synonymy with formations resulting from other processes (participles, compounding, serial compounding) is very well-tolerated and it exists as a full alternate, occasionally with lexical differentiation, side by side with derivation. Synonymy seems to be much more tolerable with stems denoting human beings and abstracta in general. However, word-formation synonymy is more stable than synonymy on any other level, as a general principle, as can be seen from the more frequent differentiation on lower levels.

[18] The instability of such differentiations is proven by the following remark found in a student newspaper (May 1986): "Viele von Euch haben es selbst erleben können, das Demokratieverhältnis der Linken; manche sogar sehr schmerzhaft."
This example has been simplified, as there was originally a differentiation on the level of situational role.

4. Conclusion

Thus far the asymmetrical relations of paradigmatic polysemy and synonymy have been portrayed as being in principle natural and tolerable, although potentially unstable (as are indeed all semantic relations). More particularly, word-formation polysemy and synonymy appear to be frequently-developing and endurable phenomena. How can the disruption of these relations be interpreted? From a psychological standpoint we must consider the value of the linguistic sign. The formal word-formation process functions as a s i g n a l, that is, the primary function of word-formation processes is precisely to "point at" the existence of some relation; the known quantities are the formal processes, the formal elements, and the semantic relations It is possible to have word formation acting on this level alone, that is, without normative restriction. Here potential synonymy can be exploited to a maximum; any one of the palette of formal candidates may be chosen to indicate a given semantic rule. We see in the modern European languages that lexical fixation plays a greater role than in previous stages: there is a noticeable preference for maximal polysemy, that is, for limiting the signal function to a fixed, known number of formal possibilities. Normative restrictions may of course be in force in a particular case or over a subset or the word-formation system.

It should now be sufficiently clear that a complex semantic structure for complex words must be recognized and that the systematic, relational level of word-formation meaning must be separated from lexical levels of meaning (lexical meaning, situational role, the semantic aspects of style). Without such a differentiated concept, it is impossible to offer an adequate description of language change in word-formation. Moreover, if the word-formation level is recognized as being the only systematic one, we can restrict our demands for predictability in diachrony and dynamic formation to this level. The discussion of p r o d u c t i v i t y as a concept concerning systematicity and rules belongs in this level alone. Likewise, word-formation synonymy and polysemy as defined here receive a special status in the context of paradigmatic polysemy and synonymy; changes in the constellations of Form-Meaning rule combinations within a paradigm are of a higher order than changes on the lower levels, and it is on the basis of t h e s e changes that the word-formation system evolves over time. One of the major forces behind the constant evolution of word-formation systems is the potential nature of word-formation polysemy, causing expansion of word-formation meaning for one form on the one hand (thus destroying "functional splits") and thus contributing eventually to increased synonymy on the other hand. The actualization of potential word-formation synonymy contributes to polysemy in a given paradigm. These interdependencies, which can be quite intricate, are the stuff that investigations of word-formation (semantics) should be based on.

Bibliography

Aronoff, M.: Word Formation in Generative Grammar. MIT Press, Cambridge, Ma. 1970.
Aronoff, M.: Word formation and lexical meaning. In: Quaderni di Semantica 5 (1984). 45–49.
Beard, R.: The Indo-European Lexicon. North Holland, Amsterdam 1981.
Beard, R.: Generative lexicalism. In: Quaderni di Semantica 5 (1984), 50–57.
Broselow, E.: Language change and theories of the lexicon. In: Papers from the 13th Regional Meeting, Chicago Linguistic Society 1977, 58–68.
Dressler, W. U.: Universalien von Agens-Wortbildungen. In: Brettschneider, G. and Lehmann, C. (eds): Festschrift für H. Seiler. Narr, Tübingen 1980, 110–113.
Duden: Das Große Wörterbuch der deutschen Sprache. Dudenverlag Mannheim 1976.
Fowler, H.: Modern English Usage. OUP, Oxford 1946
Gawełko, M.: Evolution des suffixes adjectivaux en français. Wrocław 1977.
Grimm, H. J.: Untersuchungen zur Synonymie und Synonymität im neueren Deutsch. Dissertation Leipzig 1970.
Mal'ceva, I. M.: Iz nabljudenij nad slovoobrazovaniem v jazyke XVIII v. (na meteriale odnokorennyx parallelej *-ost'*, *-stvo*, i *-ost'*, *-ie*). In: Processy formirovanija leksiki russkogo literaturnogo jazyka, Ju. S. Sorokin (ed) 1966, 259–284.
Marle, J. v.: On the Paradigmatic Dimension of Morphological Creativity. Dordrecht 1985.
Mel'čuk, I. A.: Opyt teorii lingvističeskix modelej "Smysl⇔Tekst". Moskva 1974.
Ljung, M.: English Denominal Adjectives: A Generative Study of a Group of High Frequency Denominal Adjectives in English. Göteborg 1970
Ohnheiser, I.: Untersuchungen zur Wortbildungssynonymie im nominalen Bereich (Substantiv und Adjektiv) in der russischen Gegenwartssprache. Leipzig 1974.
Pounder, A.: Systemangemessenheit in der Wortbildung am Beispiel desubstantivischer Adjektivableitung im Deutschen. Dissertation Vienna 1987.
Schupbach, R.: Lexical Specialization in Russian. Slavica, Columbus 1984.
Vinogradov, V. V. et al.: Izmenenija v slovoobrazovanii formax suščestvitel'nogo i prilagatel'nogo v russkom literaturnogo jazyke XIX v. Nauka, Moskva 1964.
Warren, B.: Semantic Patterns of Noun-Noun Compounds. Göteborg 1978.
Uluxanov, I. S.: Slovoobrazovatel'naja semantika v russkom jazyka i principy ee opisanija. Nauka, Moskva 1977.
Zwanenburg, W.: Word formation and meaning. In: Quaderni di Semantica 5 (1984)-130–142.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 197–209 (1986)

RECURSIVENESS IN WORD-FORMATION, WITH SPECIAL REGARD TO SPANISH

F. RAINER

(University of Salzburg, Austria)

1. The case for an "iteration constraint"

There seems to be fairly wide-spread agreement among students of morphology as to the existence of a general constraint to the effect that some kinds of word-formation rules or even word-formation rules in general may not apply to their own output.

As far as I can see, Uhlenbeck (1962, 428) was the first to come up with the idea "that it is not permitted to add the same affixational element twice".[1] This version of an Iteration Constraint will henceforth be referred to as "Iterated Affix Constraint".

Mayerthaler (1977, 61), Strauss and Bauer narrow the domain of the Iteration Constraint to suffixation (I shall call this version the "Iterated Suffix Constraint"). They use it to explain the ungrammaticality of **reddishish*, **syntacticalal*, **brotherhoodhood*, **dictatorshipship* (Strauss (1980, 105)), **joyfulful*, **helplessnessness*, **dukedomdom* (Bauer (1983, 92)) and **containerizeize* (ibid., 80).

Bauer further notes that

> It is not the fact of previous suffixation which blocks these forms, as the existence of *help-less-ness* proves, but the identity of the two suffixes. Note that this is not a restriction on the function of the suffixes, since two nominalization suffixes can be conjoined, as in *provis-ion-ment*, but **environ-ment-ment* is not possible.

The "Multiple Application Constraint" of Lieber (1981, 171—173), on the other hand, is even more powerful than the Iterated Affix Constraint, since it applies to word-formation rules in general. It reads as follows:

> No word formation process (. . .) can apply iteratively to its own output. (173)

In Lieber's dissertation it serves the purpose of preventing a Tagolog reduplication rule from applying more often than it should. In order to dissipate the aura of *ad hoc*-ness she adduces independent evidence from English, German and Spanish, where the ungrammaticality of **blueishishish*, **ununun-*

[1] Van Marle (1985, 62) accepts Uhlenbeck's constraint.

happy, **Mädchenchenchen*, **Vögleinleinlein*, **pequeñititito* and **muchachototote* is claimed to be attributable to the Multiple Application Constraint.

With respect to the Spanish data she adds:

> Iteration of the diminutive and augmentative affixes in Spanish can occur in only three forms: *chico*, *poco* 'small' can iterate the diminutive, presumably indefinitely, and *grande* 'large' can iterate the augmentative (cf. Harris (1979).[2] Whatever the explanation for these three forms, it is clearly the case that the augmentative and diminutive do not iterate freely (204—205).

Finally, it deserves special mention that according to Lieber (1981, 173) "this ill-formedness is certainly not a semantic phenomenon".

Botha (ms.) essentially accepts Lieber's Multiple Application Constraint and applies it to his analysis of Afrikaans reduplication in a way parallel to Lieber's treatment of Tagalog. He is not unaware of problematic data,[3] however, and concludes that these "should be analyzed in depth in a study that attempts to determine the limits of the domain within which the Multiple Application Constraint applies".

It is precisely this question that I want to take up in my contribution.

2. Recursiveness in word-formation

Since a general constraint against recursion in word-formation would automatically rule out the possibility of having identical adjacent affixes, let us first dwell somewhat on this issue.

2.1. Matthews (1972, 97—98) and Bauer (1978, 333—334) have denied the existence of recursion for the Latin word and for Germanic suffixation[4] respectively. According to Stein (1976, 226) "recursiveness does not occur in derivation", and Motsch (1979, 12) hypothesizes that there might be an upper limit ("eine grundsätzliche Grenze") to the length of complex words independent of the restrictions single affixes are subjected to.

Apart from these rather succinct observations I have found two more elaborate cases against recursion in word-formation in Reichl (1982) and in Szymanek (1982).

In Reichl's categorial study on the deadjectival abstract noun in English recursion is ruled out by subcategorizing English nouns in abstract and non-

[2] Harris' paper "has not been published, nor will it be. Rather, it has met the quiet death that it deserved" (personal communication).

[3] Polish iterated diminutives in *-ecz-ecz-ek*, pointed out to Botha by Dressler, are ungrammatical according to Szymanek (1982, 175—177). They rather conform to than disprove the Multiple Application Constraint.

[4] Bauer refrains from generalizing the Germanic case because "Postal reports cases of cyclical word-formation in Mohawk". According to Reich (1969, 834), however, Postal's data are unreliable.

abstract ones, English adjectives in derived and nonderived ones, and by further preventing the category "abstract noun" from turning up to the right of the category "derived adjective". All he achieves by this move is the possibility to state, case by case, whether there is recursion or not. His treatise does not embody, however, any theory about which word-formation rules may be recursive and which not, and why this is so.

Szymanek presents an interesting set of data from Polish, where the deadjectival nominal suffix *-ość* and the denominal adjectival suffix *-ow(y)* can feed one another (cf. (1) and (2)):

(1) *cał* → *cał-ość* → *cał-ość-owy* → *cał-ość-ow-ość* → **cał-ość-ow-ość-owy*
(2) *lud* → *lud-owy* → *lud-ow-ość* → *lud-ow-ość-owy* **lud-ow-ość-ow-ość*

Szymanek essentially accounts for these and similar data by means of a filter forbidding "two phonetically identical (groups of) suffixes, in a cyclically derived word" (182). Limited recursion in word-formation thus seems to be viewed as the result of the interaction of the morphological and the phonological components of a grammar, though no explanation is given as to why a similar phonological filter should exist.

Szymanek's filter is an extension of Dressler's (1977) "Hapological Constraint", and so is Menn/Mac Whinney's (1984) "Repeated Morph Constraint", which reads as follows:

*XY,-where X and Y are adjacent surface strings such that both could be interpreted as manifesting the same underlying morpheme through regular phonological rules, and where either

(a) X and Y are both affixes, or
(b) either X or Y is an affix, and the other is a (proper subpart of a) stem.[5]

At first sight, the Iterated Affix Constraint might seem to be a straightforward consequence of this constraint. If I am nevertheless not inclined to adopt this hypothesis, it is mainly because Menn/Mac Whinney's constraint applies only very sporadically, under conditions which are not yet fully elucidated. By and large it is restricted to accidental morpheme repetition. This condition, however, is never met in the application of a word-formation rule to its own output.[6]

Summing up this section I feel prompted to conclude that none of the anti-recursionists has presented really convincing arguments for his standpoint.

[5] Interestingly Dressler (1977, 44) already noted that "the haplological constraint can be violated, if a suffix is repeated for intensification".

[6] It is probably unnecessary to postulate an *ad hoc* phonological constraint in order to rule out the iteration of the Dutch prefixes *be-*, *ver-* and *ont-*, as Booij does, if one takes into account the semantics of these prefixes.

2.2. The advocates of recursiveness in word-formation have adduced two kinds of arguments for their viewpoint: on the one hand, they have pointed to some clear instances of recursive word-formation rules in natural languages, and on the other, they have tried to get rid of the problematic cases by assigning their unacceptability to extragrammatical factors.

The originator of this type of argument repeated again and again in the literature (cf. Chomsky (1970, 212), Booij (1977, 153), Guerssel (1983, 241) with explicit reference to their source, and Lightner (1975, 619), Dell (1979, 194—195), Corbin (1980, 84—91), Carden (1983, 540—541), without reference) seems to have been Chapin, who writes:

> With cyclic patterns (. . .) one might expect the possibility of recursion. This expectation is born out (. . .). Cases in point are derivative sequences like *organize, organization, organizational, organizationalize, organizationalization*, etc.; *physical, physicalist, physicalistic, physicalistical, physicalisticalist*, etc. Each new derivative receives an increasingly precise semantic interpretation. By the end of the second cycle, the derivative is so narrowly precise as to be totally useless, and so internally complex as to be difficult to understand. This does not make it an impossible form, however, any more than a sentence with multiple self-embeddings, which would never be used in an actual utterance, is thereby ungrammatical. (1970, 60)

Thus, according to Chapin, the unacceptability of certain recursive formations should not be traced to their morphological ill-formedness, but rather to pragmatic and cognitive factors. Though such an explanation, in the absence of explicit theories of pragmatics and cognition, might be considered by the malevolent as an immunizing strategy, it seems plausible enough to me to adopt it in the subsequent discussion.

3. Arguments against iteration constraints

I shall now turn to the critical appraisal of the different Iteration Constraints presented in **1.**

3.1. Cave restrictiones

Let us begin with some obvious, but nevertheless necessary, remarks.

It goes without saying that a word-formation rule can only apply to its own output if the output qualifies as an input to this rule (cf. Stein (1977, 225), Plank (1981, 127)). As the reader may already have noticed, not all the morphologists quoted in 1. are in accordance with this fundamental requirement.

3.1.1. Category-changing affixes can obviously never be iterated (cf. Lieber (1981, 72), Mayerthaler (1981, 118)). Thus Bauer's **joyfulful*, **helplessnessness* and **containerizeize* turn out to be pseudoproblens.

3.1.2. If we allow word-formation rules to place semantic restrictions on their base, we can equally forget about Bauer's **dukedomdom* and Strauss' **brotherhoodhood*, **dictatorshipship*.

But semantic restrictions may even be much subtler.

The English prefix *un-* is commonly held not to attach to negative bases: if this is true, it would suffice to explain the unacceptability of Lieber's **ununhappy*, Bauer's *ununlikely* (1983, 69) and Stein's **ununfair*. The same explanation is available for French **inillégal* and **anamoral* (Corbin (1980, 83—84)).

3.1.3. Morphological restrictions may also come into play.

Since the English suffix *-ful* seems to be restricted essentially to monomorphemic bases Reichl's (1982, 19—20) **pitifulnessfulnessful* is not particularly well chosen to demonstrate the limited recursiveness of word-formation rules.

3.1.4. A phonological or euphonic restriction is often said to be responsible for the oddness of **fishish* and **bitchish* (cf. Neuhaus (1971)). If we accept this restriction, we lose two more alleged Iterated Affix Constraint explananda: Strauss' **reddishish* and Lieber's **blueishishish*.

In this case, however, another factor might conspire with the euphonic restriction: "approximative" affixes like *-ish* do not seem to be iterable in any natural language (cf. German **grünlichlich*, Dutch **groenigig*, French **verdâtrâtre*, Spanish **negruzcuzco*, Italian **giallognolognolo*, etc.; *pace* West Friesian) van Marle (1984, 161, footnote 9), be it for pragmatic reasons, as Plank (1981, 128) suspects, or for cognitive ones ('having the quality of being near to the quality of being near to the quality of being green' might be argued not to be a well-formed concept).

3.2. Iterated affixes

We have seen thus far that many of the cases adduced as alleged Iteration Constraint explananda turn out to be pseudoproblems on a closer inspection, which considerably weakens the case for the need for such a constraint.

I shall now further call this view into doubt by presenting counterexamples, i.e. cases of iterated affixes.

3.2.1. Iterated prefixes

As we have seen, Bauer, Mayerthaler and Strauss restrict the domain of the Iteration Constraint to suffixation, a wise move, since repeated prefixes seem to be quite common even in the more familiar European languages.

Booij (1977, 154) contains a long list of Dutch formations such as *ultra-ultramodern, meta-metataal, para-paranormaal, ex-exechtgenoot, anti-anti-raket-raket-raket, bet-betovergrootvader, over-overgrootvader, achter-achterkleinzoon, sub-subgroep, super-supermooi, her-heroverweeg.* <6>

Corbin (1980: 83—84) mentions the following iterable prefixes from French: *re-, anti-, contre-, super-, archi-, ultra-, pré-, avant-, post-, après-.*

In the Spanish of San Luis and Rosario (Argentina) the intensive prefixes *re-* and *réquete-* are reported to be commonly iterated (Vidal de Battini (1949, 215), Donni de Mirande (1968, 75)). And in a cartoon by Juan Ballesta (Cambio 16, 20—5—85, 3) a postmodernist philosophizes: *Es lógico que si ayer éramos posmodernos, hoy somos posposmodernos y mañana seremos pospospos modernos.*

3.2.2. Iterated suffixes in Spanish

Examples of iterated suffixes seem less easy to come by.

This, however, is what we should expect if we consider that in the Indoeuropean languages suffixation is typically category-changing, while category-constancy is often claimed to be a defining property of prefixation. This means that by virtue of this general characteristic of Indoeuropean derivational systems the number of potential candidates for iteration will by its very nature be extremely scarce among suffixes.

There nevertheless do exist some specimens, as I intend to show in the following discussion of the situation in Spanish.[7]

3.2.2.1. The first attested iterated suffix goes back to the 16th century: *son muy poquitos y aun poquititos los sabios* (Gon, 205—207).[8] As González Ollé rightly observes, the absence of iterated forms in the early literature is probably due to the highly colloquial flavour of such formations, not to their non-existence.

The first explicit reference to iterated suffixes I have found in Correas' grammar from 1627: *tanbien en duplicarlas* (sc. diminutives (and augmentatives?)) *ai mucha libertad* (146). Among the examples we find *tantitito, tamarrinino, tamañinino.*

[7] The cases of iterated suffixes reported in the literature are extremely small in number.

Afrikaans is sometimes claimed to have iterated diminutives (Schultink (1974)), but the situation is not quite so clear (Botha (ms.: 186)).

In Dressler (1977, 44) one can find the Latin iterated diminutive *homullulus.*

Van Marle (1985, 107) reports playful iteration of the compararive suffix in Dutch (*leuk-er-der-der* 'much funnier' from *leuk* 'funny').

West-Friesian (van Marle (1985, 161)) seems to tolerate the iteration of the approximative suffix *-ig.*

Iterated causatives are reported to exist in Turkish and Quechua (Plank (1981, 263), in Caribian (Hoff (1981)) and, maybe, in Japanese (Chomsky (1982, 96)).

[8] In oder to save space, the titles from the subsection "Spanish" of the bibliography will be referred to by quoting only the first three letters of the name of the author. In the case of the two Sanchezes the first letter of the second name is added for the sake of clarity.

3.2.2.2. In the following exposition of present-day usage I shall order the material according to suffixes starting with the diminutive suffixes, which are the most common, and then turning to augmentative and intensive ones.

The material has mainly been gathered from articles on word-formation and from monographs on colloquial and dialectal usage in Spain and Latin America. In fact, the standard language seems to be much less prone to iteration than the "lower" varieties.

3.2.2.2.1. The iteration of the diminutive suffix *-ito* seemingly occurs in all colloquial varieties.

I have explicit attestations for Andalusia (*chiquitito*, Bei, 294), Mexico (*toditito, la puritita verdá*, Wal, 13, 16; *chiquitito*, Ran, 107; *poquitito, toditito, oritita, orititita*, Boy, 127; *ahoritita, toditito, chiquitito, chiquititito*, Pra: 90, 78, 44; *lo primeritito que vi*, Gaa: 586), rural Panama (*chiquitito, chiquititito, poquitito, poquititito*, Rob: 71), Ecuador (*solitito*, Wal: 15), Peru (*chiquitito*, SanD: 24), the dialects of San Luis and Rosario (Argentina), where the process is reported to be very productive (*chiquito, chiquitito, chiquititito, chiquitititito, rechiquititito, réquetechiquititito, réqueterrechiquititito, cerquita, cerquitita, cerquititita*, etc., Vid: 350; *chiquitito, cerquitita*, Don: 91), and for Chile (*poquitito, chiquitito, toitito, naitita*, Len: 196; *lueguitito, pedacitito*, Oro: 276, 277).

In Costa Rica (Zam: 545), Santo Domingo (Hen, 194) and Columbia (Fon, 558) instead of pure iteration of *-ito* we find sequences of *-ito* and *-ico*: *chiquitiquitiquitico* (Fon). This process, by the way, is already present Correas' grammar.

3.2.2.2.2. Iteration of the diminutive suffix *-ino* is attested for the Leonese dialect of Cespedosa de Tormes (SanS: 165—166). This geographical limitation comes as no surprise if we consider that *-ino* itself is essentially limited to this dialect.

> En los adjectivos se añade *-ino* a los que por si mismos expresan idea de pequeñez, como *delgado, bajo, corto, ralo*, etc., y es frecuente duplicar y aun triplicar el sufijo para acentuar el sentido diminutivo; rara vez se oye *chiquino, cortino* o *bajino*, sino más bien *chiquinino, bajinino, certinino*, y en vez de *delgadino* se dice *delga-in-ino, delga-in-in-ino* o *delga-irr-in-ino*. Lo mismo ocurre con los adverbios: *cerca*, de que se forman *cerqu-in-ina, poqu-in-ino*; de *pizca* se dice una *pizqu-in-ina* 'un momentín'; de *puñado*, un *puña-in-in*, y *unos cuantininos*, de la frase 'unos cuantos'.

3.2.2.2.3. The Spanish of Chile has an affective diminutive suffix *-icho*, which can also be iterated: *poquichicho, chiquichicho, toichicho, naichicha* (Len, 196); *al tirichicho* (Rab, 245); *lueguichicho* (Oro, 276). According to Rabanales (1958, 245) it is "más propio del habla rural que de la urbana".

3.2.2.2.4. It deserves special mention that the diminutive suffix *-illo* never seems to be iterated.

3.2.2.2.5. Iteration of the augmentative suffix *-ote* is attested for the Leonese dialect of Cespedosa de Tormes (*arribotota, abajotote,* SanS: 168) and several Latin-American countries. Unfortunately, most of the Latin-American sources only refer to one example, i.e. *grandotote* (Ran, 107, Boy, 131, Gaa 590, for Mexico; Rob, 82, for rural Panama; Hen, 191, for Santo Domingo). Only Flo, 54 has *arribotota* and *es mucho entierrotote,* glossed as 'con muchas flores y concurrencia de numeroso público', for Santander (Columbia).

3.2.2.2.6. The augmentative suffix *-ōn* is also iterated in Santander (Columbia): *tiendonón, le dan unos malonones, el pilonón de plata, mazorconón,* Flo: 53.

3.2.2.2.7. Iteration of the augmentative-intensive suffix *-azo* is attested to occur in San Luis (Argentina), while Oro, 286 explicitly denies the existence of such iterated forms for Chile.

> Es común la reduplicación del sufijo, que se convierte en *-azazo, -azaza*: *bocazaza, amigazazo, ladronazazo, lejazazo, fierazazo.* (Vid, 371)

3.2.2.2.8. Intensive *-ísimo* is reported to be — partially! — iterated as *-isísimo* by several authors. Bei, 285 has "playful" *requetemonisísima,* Boy, 134 *riquisísimo,* Flo, 51 *muchisísimo.* In rural Panama "intensification is indicated through repetition of *-is-*, although speakers do not overwork this device: *hermosisísima, malisísimo, riquisísimo*" (Rob, 83). The most detailed description can again be found in Vidal de Battini's monograph:

> Es frecuente reduplicar la sílaba *-si-* del sufijo, para aumentar su valor intensivo. Así se dice: *malísimo, malisísimo, malisisísimo*; también con los prefijos *re-* y *réquete*: *remalisísimo, requeterremalisísimo.* (Vid, 373)

3.2.2.3. I would like to close this section on iterated suffixes in Spanish with the following remarks:

a) Due to the scarcity of information available about wordformation in colloquial and dialectal varieties of Spanish my synopsis is necessarily incomplete, especially as far as the actual productivity and geographical extension of the processes is concerned.

b) The synopsis nevertheless warrants the conclusion that Lieber's claims about the non-occurrence of iteration in Spanish were, to say the least, premature. Iterated suffixes do occur in Spanish, though not equally frequently in all of its varieties. As I have already pointed out, they are typical of "lower" varieties, while educated and literary usage tends to reject them.

4. Do we really need an iteration constraint?

4.1. I have opened this contribution by quoting some weighty voices advocating the necessity of some kind of Iteration Constraint in order to avoid the formation of lots of unacceptable forms.

The evidence from Dutch, French and Spanish prefixation which, by the way, could easily be enriched by similar data from other languages, suffices, I think, to disprove both the Multiple Application Constraint and the Iterated Affix Constraint. It seems to be the case that any rule of prefixation whose output qualifies as an input to this same rule may be iterated.

It should not be denied that besides the many positive instances for this claim there are some more recalcitrant examples, like *ex-ex-ex-husband*, but I think that the oddness of similar cases may normally be traced to extra-morphological factors after the manner of Chapin (1970). In the example under discussion one should take into account that such a formation is on the one hand extremely hard to parse ('a person who ceased to cease to cease to be a husband') and on the other hand quite useless since very few possible denotata for it exist in our real world, if any.

4.2. The prefixational data obviously have no bearing on the Iterated Suffix Constraint.

The Spanish data, however, as well as the additional evidence allued to in footnote 7, seem to disprove even this less powerful version.

Apart from this empirical argument one might also criticize that the Iterated Suffix Constraint, as it stands, is devoid of any explanatory value, since it leaves in the dark which distinguishing properties of suffixes as opposed to prefixes might be responsible for their different behaviour with respect to iteration.

Even a fierce defender of the Iterated Suffix Constraint will subscribe to this latter critique and he may even be willing to admit that the Spanish data are in some way problematic. But he will also not fail to retort that it is then up to his critics to explain the oddness of, say, German **Mädchenchenchen*, **Vögleinleinlein*, English **streamletlet*, and the like.

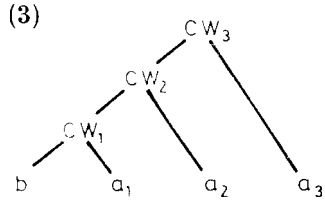
One might try to find independent reasons for the oddness of each of the above examples.[9] But I think this is not the right step to take. Instead, I will try to give a more principled and unitary account of these facts.

[9] P. Hummer rightly has pointed out to me that my semantic account does not predict the ungrammaticality of German **Mädchenchen*, where the first *-chen* is no longer a diminutive suffix synchronically. My guess about what is wrong with *-(ch)enchen* goes in the following direction: since German words in *-en* lose this "ending" in derivation (cf. *Garten/Gärtchen*, etc.), the regular diminutive of *Mädchen* would be **Mädchchen*, which is clearly not a phonotactically wellformed word in German.

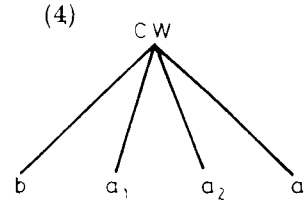
4.3. This account rests on the following — more or less uncontroversial[10] — assumptions about the structure and interpretation of morphologically complex words.

a) Derivational word-formation rules have binary structure. A complex word (CW) consisting of a base (b) and three affixes (a_1, a_2, a_3) has the structure of (3), not of (4):

(3)



(4)



b) The semantic interpretation of complex words obeys the principle of compositionality, i.e. the meaning of CW_1 is the result of applying (the function) a_1, to b, the meaning of CW_2 of applying a_2 to CW_1 or, more generally, the meaning of CW_n is the result of applying a_n to CW_{n-1}.

4.4. With these two prerequisites in mind we can now come back to our problem.

A word like *meta-meta-meta-language* is completely in accordance with the principles of binarity and compositionality just expounded. Its meaning is, as it should be, 'a language used to talk about a language used to talk about a language used to talk about a language'. In the same way German *vorvorgestern* or Spanish *anteanteayer* mean 'on the day before the day before today', while German *überübermorgen* or French *après-après-demain* mean 'on the day after the day after today'.

Now, what about diminutives, augmentatives, comparatives, intensives, and the like?

According to the principles of binarity and compositionality German **Häuschenchen*, if it existed, would mean 'a small specimen of the class of small specimens of the class of houses' (assuming a mainly quantifying function of the German diminutive), German *urururalt* would mean 'having to a considerable extent the quality of having to a considerable extent the quality of having to a considerable extent the quality of being old', while Dutch *leukerderder* would mean 'having to a greater extent the quality of having to a greater extent the quality of having to a greater extent the quality of being funny'.

But this is not the way we use and understand such iterated words (though perfectly "logical", they seem to run counter to some arcane principle

[10] For a more elaborate view cf. Pesetsky (1985), for an anticompositional stand Plank (1981).

of conceptuel wellformedness). The iteration in such cases does not result in higher-order qualities, rather it is simply tantamount to an intensification of the meaning of the affix. Similar considerations probably hold for syntactic phrases like *very very very old.*

How do we get this intensive interpretation in a principled way? My guess is that in such cases the following interpretational principle (let us call it "Iterated Intensifiable Affix Principle") comes into play:

> The iteration of an intensifiable affix may be interpreted as an intensification of that affix.

This is but an instance of the general iconic principle 'more substance = more meaning', which, applied to intensifiable morphemes, gives 'more substance = higher intensity'. This same principle is probably responsible for the well-known fact that reduplication most commonly serves an intensifying function in the languages of the world.

4.5. My account thus runs as follows:

a) We do not need any kind of Iteration Constraint. On the contrary, any word-formation rule whose output qualifies as an input to this same rule may be iterated. Of course, there may be hindering factors of an extra-morphological nature.

b) Iterated diminutive, augmentative, intensive, comparative, approximative and possibly some other affixes cannot be regularly = compositionally interpreted due to some — admittedly somewhat arcane — principle of conceptual well-formedness forbidding higher-order qualitities.

c) If the affixes mentioned in b) are nevertheless iterated, the Iterated Intensifiable Affix Principle will assign an intensive meaning to this iteration.

4.6. To conclude, I will briefly discuss some of the predictions and positive consequences of my account.

a) Several compositionally iterable affixes may be subjected to the Iterated Intensifiable Affix Principle as well. Thus a formation like *überüber-übermorgen* either refers to 'the day after the day after the day after today' (compositional reading) or more vaguely to some day in the remote future (intensive reading). Similarly, *metametametalanguage* may be used to refer to some language very high up in the hierarchy of metalanguages. Such semantic ambiguities are to be expected under my account.

b) Since the account is intended to hold for word-formation in general, it delimits the set of possible iterable affixes and their respective interpretations for all languages of the world. It is thus easily testable.

c) This account does not predict why, for example, iterated diminutives occur in Spanish but not in German or English.

It may, however, help to see why it is with diminutives and the other categories subjected to the Iterated Intensifiable Affix Principle thet we find particularly significant cross-linguistic differences. It seems to be the case that it is the "primitiveness" of this interpretational principle which is responsible for the limitation to "lower" language varieties we have encountered for such categories.

The Spanish data I have collected agree with my explanation, which of course comes as no surprise, since it was designed to fit them. Whether the remaining 3000 languages of the world will turn out to be equally well-behaved remains to be seen.

References

a) General linguistic literature:

Bauer, L.: Productivity in word-formation. In: Gregersen, K. (ed.): Papers from the Fourth Scandinavian Conference of Linguistics. Odense 1978, 333–337.

Bauer, L.: English Word-formation. Cambridge 1983.

Booij, G.: Dutch Morphology. A Study of Word Formation in Generative Grammar. Lisse 1977.

Boons, J.-P.: Sceller un piton dans le mur. Desceller un piton du mur. Pour une syntaxe de la préfixation négative. In: Langue Française 62, 1984, 95–128.

Botha, R.: Morphological from, semantic interpretation, conceptual structure: a Galilean analysis of Afrikaans reduplication. Manuscript.

Carden, G.: The non-finite-state-ness of the word formation component. In: Linguistic Inquiry 14, 1983, 537–541.

Chapin, P.: On Affixation in English. In: Bierwisch, M. and Heidolph, K. E. (eds): Progress in Linguistics. The Hague, 1970, 51–63.

Chomsky, N.: (1970) Remarks on Nominalization. In: Jacob, R. and Rosenbaum, P. (eds): Readings in English Transformational Grammar. Waltham, Mass. 1970, 184–221.

Chomsky, N.: The Generative Enterprise. A discussion with Riny Hugbregts and Henk van Riemsdijk. Dordrecht 1982.

Corbin, D.: (1980) Compétence lexicale et compétence syntaxique. In: Modèles linguistiques II. 2, 1980, 52–138.

Dell, F.: (1979) La morphologie dérivationnelle du français et l'organisation de la composante lexicale en grammaire générative. In: Revue Romane 14, 1979, 185–216.

Dressler, W.: (1979) Phono-morphological Dissimilation. In: Dressler, W. and Pfeiffer, O. (eds): Phonologica 1976. Innsbruck 1977, 41–48.

Guerssel, M.: A lexical approach to word formation in English. In: Linguistic Analysis 12, 1983, 183–243.

Harris, J.: Explorations in Spanish Word-Structure: An X-bar Theory. Manuscript 1979.

Hoff, B.: De Caraibische causatief als productief en iteratief procédé. In: Forum der Letteren 22, 1981, 148–158.

Lieber, R.: On the Organization of the Lexicon. Bloomington 1981.

Lightner, T.: The role of derivational morphology in generative grammar. In: Language 51, 1975, 617–638.

Marle, J. van: On the Paradigmatic Dimension of Morphological Creativity. Dordrecht 1985.

Matthews, P.: Inflectional Morphology. A Theoretical Study Based on Aspects of Latin Verb Conjugation. Cambridge 1972.

Mayerthaler, W.: Studien zur theoretischen und zur französischen Morphologie. Tübingen 1977.

Mayerthaler, W.: Morphologische Natürlichkeit. Wiesbaden 1981.

Menn, L. and Mac Whinney, D.: The Repeated Morph Constraint: toward an explanation. In: Language 60, 1984, 519–541.

Motsch, W.: Zum Status von Wortbildungsregularitäten. In: DRLAV-Papier 20, 1979, 1—40 (Paris).
Neuhaus, H.: Beschränkungen in der Grammatik der Wortableitungen im Englischen. Saarbrücken 1971.
Pesetsky, D.: Morphology and Logical Form. In: Linguistic Inquiry 16, 1985, 193—246.
Plank, F.: Morphologische (Ir-)Regularitäten. Aspekte der Wortstrukturtheorie. Tübingen 1981.
Reich, P.: The finiteness of natural language. In: Language 45, 1969, 831—843.
Reichl, K.: Categorial Grammar and Word-Formation: The Deadjectival Abstract Noun in English. Tübingen 1982.
Schultink, H.: Dubbele diminutieven in het Afrikaans en transformationeel-generatieve taalbeschrijving. In: Taalkunde — 'n lewe. Studies opgedra aan prof. dr. W Kempen by geleegenheit van sy 65ste verjaardag. Kaapstad 1974.
Stein, G.: Semi-Productive Lexical Rules: A Note on *-ed-* Adjectives. In: Journal of English Linguistics 10, 1976, 30—33.
Strauss, S.: How abstract is English morphology? In: Glossa 14, 1980, 89—112.
Szymanek, B.: Cyclic Word Formation Rules and the Haplological Constraint. In: Studia Anglica Posnaniensia 14, 1982, 175—84.
Uhlenbeck, E.: Limitations of Morphological Processes: Some Preliminary Remarks. In: Lingua 11, 1962, 426—432.

b) Spanish:

Beinhauer, W.: El español coloquial. Madrid 1978.
Boyd-Bowman, P.: El habla de Guanajuato. México 1960.
Correas, G.: Arte kastellana. 1627. Ed. by M. Toboada Cid. Santiago de Compostela 1984.
Donni de Mirande, N.: El español hablado en Rosario. Rosario 1968.
Flórez, L.: El español hablado en Santander. Bogotá 1965.
Fontanella, M.: El diminutivo en Bogotá. In: Thesaurus 17, 1962, 556—573.
Gaarder Bruce, A.: Los llamados diminutivos y aumentativos en el español de México. In: Publications of the Modern Language Association of America 81, 1966, 585—595.
González Ollé, F.: Los sufijos diminutivos en castellano medieval. Madrid 1962.
Henríquez Ureña, P.: El español en Santo Domingo. Buenos Aires 1940.
Lenz, R.: La oración y sus partes. Madrid 1920.
Oroz, R.: La lengua castellana en Chile. Santiago de Chile 1966.
Prado Garduño, G.: El uso del diminutivo en el habla familiar de la ciudad de México. México 1964.
Rabanales, A.: Recursos lingüísticos en el español de Chile, de expresión de la afectividad. In: Boletin de Filología (Santiago de Chile) 10, 1958, 205—302.
Ranson, H.: Diminutivos, Aumentativos, Despectivos. In: Hispania 37, 1954, 406—408.
Robe, S.: The Spanish of Rural Panama. Major Dialectal Features. Berkeley 1960.
Sánchez Diaz, A.: Sobre el diminutivo en la Sierra Peruana. In: Lenguaje y Ciencia (Trujillo, Peru) 24, 1966, 23—26.
Sánchez Sevilla, P.: El habla de Cespedosa de Tormes (en el límite de Salamanca y Avila). In: Revista de Filología Española 15, 1928, 131—172.
Vidal de Battini, B.: El habla rural de San Luis. Parte I. Buenos Aires 1947.
Walsh, D.: Spanish Diminutives. In: Hispania 27, 1944, 11—20.
Zamora Elizondo, H.: Los diminutivos en Costa Rica. In: Thesaurus 1, 1945, 541—546.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 211–219 (1986)

TRANSPARENT HEAD, INHERITANCE AND THE NORMAL FORM*

J. TOMAN

(University of Michigan, USA)

In the following talk I would like to address two points. The first deals with three sets of phenomena which have hitherto been treated separately in the literature on word-formation. I would like to suggest that a significant generalization is missed unless these phenomena are all treated under the heading of inheritance across a transparent head. Whereas this part of my talk is essentially descriptive in that it merely establishes an empirical generalization covering a certain class of facts, the second part should perhaps be more interesting. In it I will speculate about the nature of head-transparency and the representation of transparent heads at the level of Logical Form.

1. Three cases of inheritance

1.1. The data to be discussed are relatively simple. As is well known from various theoretical frameworks, the syntax of an argument-taking item such as a verb or an adjective is generally not preserved throughout morphological operations, whereas the argument structure may often remain unaffected. To take a familiar example: the syntax of a deverbal noun is not identical to the syntax of the verb from which the noun is derived, but the argument structure of a noun such as an action nominal is in an intuitive sense on a par with that of the verb. There have been numerous accounts of this phenomenon in the literature. The one that currently appears most attractive is an account in terms of inheritance, or, percolation, of properties of subparts of a word to the whole (complex) word. There have been a number of fairly detailed discussions about certain aspects of inheritance, yet there is one aspect which has not played a very important role in recent discussions. Note that it is a typical state of affairs that affixes — and I will restrict the discussion to suffixes here — do not block inheritance. Thus nominalizing suffixes typically do not

* This is the text of a talk given at the International Conference on Word Formation in Veszprém, Hungary, May 1—4, 1986

Acta Linguistica Academiae Scientiarum Hungaricae 36, 1986
Akadémiai Kiadó, Budapest

prevent arguments of an argument-taking item from being mapped onto the whole NP. In other words, an affixal head generally does not impose a barrier on inheritance. Let us call such a head a transparent head and exemplify it with such suffixes as German *ung* (for action nominals) or *bar* (for deverbal adjectives such as *lesbar* 'readable'). English *able* in adjectives such as *readable* would in fact be a good example too. They all can be subsumed under the following descriptive schema:

(1) $[[X_{\text{argument-taking}}]\ [\text{head}_{\text{transparent}}]]$

Without going into a detailed factual discussion at this point, we note that this state of affairs sharply contrasts with the realization of argument structure in compounds. Given a verb-noun compound, the argument structure of the verb will typically be blocked by the noun in the head position. To be more precise, the noun in the head position will either be an argument of the verb such as in *Eßapfel* 'eating-apple' or it will be unrelated to the verb in terms of (primary) argumenthood: *Lesestunde* 'readinghour'. But in none of these examples will argument inheritance extend across the head, i.e., examples such as:

(2) (a) **Eßapfel durch Kinder*
'eat-apple by children'
(b) **Lesestunde von Büchern durch Kinder*
'reading-hour of books by children'

are on the verge of unintelligibility. We may thus say that full nominal heads are non-transparent and set up the following schema for examples listed in (2):

(3) $[[X_{\text{argument-taking}}]\ [\text{head}_{\text{non-transparent}}]]$

It would now seem that the contrast between (1) and (3) correlates with the distinction between suffixes and nonsuffixes. Recall that items that turned out to be transparent were suffixes, whereas those which were not, were full nouns. However, I would prefer not to reduce the difference between these two classes of facts to the distinction between suffixes and non-suffixes so quickly. The reason for not doing so is motivated by a type of data which is in some sense marginal; nevertheless, it often happens that fine distinctions and conceptual clarifications are obtained through untypical cases, and this, I believe, is also the case here. Note the following examples from German which do not fit the above generalization according to which heads of compounds are not transparent:

(4) (a) *der Beschleunigungsgrad der Partikeln* (cf. *etwas beschleunigen*)
acceleration-degree of particles
(b) *die Vorbereitungszeit auf den Flug* (cf. *sich auf etwas vorbereiten*)
preparation-time on the flight
'preparation-period for the flight'
(c) *die Wachstumsgeschwindigkeit der Pflanzen* (from Höhle 1982)
growth-speed of plants
'growth-rate of plants'

The first example is particularly telling: we see that the argument structure of *beschleunigen* 'accelerate', and, derivatively, of *Beschleunigung* 'acceleration', is preserved outside the compound, i.e., across a compound head. At the same time, we however observe that *Grad* 'degree' is a somewhat peculiar noun in that it is essentially a bound nominal in the given meaning:

(5) **Wir sprechen über den Grad.*[1]
we talk about the degree'

Clearly, what we are dealing with is a kind of measure noun, i.e., a quantifier expression, which requires a certain kind of syntactic complement.

As for other examples listed in (4), a clear generalization is not easy to formulate although one has a definite intuition about them: on the whole, we are dealing with nouns which denote concepts that are very general and unspecific in nature, not with names of particular objects or concrete subspecies of general concepts. This must be the relevant direction of generalization since contrasts of the following sort can be constructed:

(6) (a) **die Vorbereitungshalle auf den Flug* (cf. 4b)
the preparation-hall for the flight
(b) **die Wachstumsstudie der Pflanzen* (cf. 4c)
growth-study of plants

We may add at this point that these contrasts are relative. Whereas not all native speakers of German may be quite happy about (4) (b), (c), they will not hesitate to indicate a contrast between these examples and (6) (a), (b). Similar relative contrasts may also be reproduced for English, as pointed out to me by Jean Boase-Beier:

(7) (a) *a combination process of quicksilver and gold*
(b) **a combination procedure of quicksilver and gold*

[1] The example might be grammatical under certain discourse conditions involving contrast etc. I take it that this fact is not important in the context given

This type of example is thus of much use in helping us to identify the relevant property of transparent heads. I would like to claim that head-transparency is not constrained to suffixes: true, suffixes are generally transparent, but nouns, if sufficiently "abstract" (or, "empty"), can behave in the same way as transparent suffixes with respect to argument inheritance. The relevant property is thus not "to be a suffix" but "to be semantically light," whatever this may mean in formal terms.[2]

1.2. Let us now consider a set of data which, especially in German literature, has fairly often been commented upon, especially by prescriptive grammarians. The question is whether it is permissible to modify a compound, typically an NN-compound, by an adjective in such a manner that the adjective relates to the noun in the non-head position of the compound, i.e., not to the head of its projection.[3] Examples quoted in the literature to show that this is impossible and should thus not be done range from grotesque formations, which may well in fact be linguists' inventions, such as:

(8) **dreiköpfiger Familienvater*
'three-headed family-father'

in which the intended reading is 'a father of a three-head family' (and obviously not 'a three-headed family-father') to relatively acceptable examples such as:

(9) *psychologische Beratungsstelle* (from Bergmann 1980)
psychological counselling-board
'board for psychological counselling'

in which the adjective plausibly relates to the noun in the non-head position.

Incidentally, a number of relatively acceptable examples show the following characteristic property: both the head noun and the noun in the modifier position can meaningfully be combined with the adjective.

(10) *deutsche Literaturwissenschaft*
German literary-science

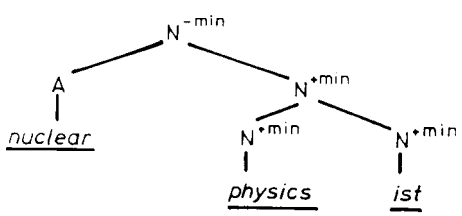
the intended reading being 'the study of German literature' rather than 'the German study of literature'. How the intended reading is exactly obtained in these cases is not very clear: I suspect that some kind of marginal non-core

[2] For further examples and discussion see Höhle (1982) and Toman (1983)

[3] See Bergmann (1980) for a summary of these opinions and a classification of basic facts. A comment on Bergmann is Sandberg (1984)

process is involved. Really good examples in which the relation can only be established between the adjective and the modifying noun are not easy to find. But note that this is exactly what we would expect if the descriptive generalization in terms of head-transparency were correct. Given this generalization, it should not be possible to link the adjective with the noun in the non-head position across a non-transparent head. In other words, it seems that the examples of adjectival modification just discussed should be treated in the same manner as examples of argument inheritance.

The prediction made by the transparency-generalization also extends to instances in which the link between an adjective and a compound-internal noun should be well-formed. This should be so with transparent heads. And indeed, we see that prototypic transparent heads, suffixes, generally permit of the link between the adjective and the noun: recall the set of examples discussed under the heading of "relatedness paradoxes" by E. Williams (1981). The relevant examples include English *generative grammarian* as well as its German counterpart *generativer Grammatiker*. It would seem that no particular device beyond inheritance of lexical properties across a transparent head should be necessary to derive the cases E. Williams brought into the discussion.[4]



1.3. To conclude the first section, let us briefly note one more set of data which fits here quite well and which must in fact also be discussed in the context of head transparency. Note that "manner adjectives" may modify deverbal nouns:

(11) (a) *starker Raucher*
strong smoker
'heavy smoker'
(b) *plötzliche Abfahrt*
'sudden departure'

If we understand this modification as a result of percolation of verbal properties to the whole nominal, this is expected, because suffixes forming action nominals

[4] Without going into details at this point, *nuclear physicist* would be derived from: by percolation of the obligatory "idiom" subcategorization of *physics* to *physicist*. I take it that on the required reading *physics* subcategorizes for the adjective *nuclear*

are generally transparent. On the other hand, the heads in the following VN-compounds are not:

(12) (a) **langsamer Eßapfel* (intended reading: 'apple one eats slowly')
slow eating-apple
(b) **plötzliche Schrumpfleber* (intended reading: 'liver that suddenly shrinks')

These latter examples are comparable to those in (2) and (6): a non-transparent head is involved. Consequently, inheritance of lexical properties from the non-head position is blocked.

To summarize the first part of this paper: we see that a unified account of argument-inheritance of the type exemplified by *Beschleunigungsgrad der Partikeln* 'acceleration-degree of particles' as well as by adjectival modifications such as *generativer Grammatiker* 'generative grammarian' or *plötzliche Abfahrt* 'sudden departure' is necessary, since all involve reference to the same concept, namely that of head-transparency. This means, among other things, that transparency involves the entire range of lexical properties, not only argument inheritance.

2. Some consequences for the theory

So far I have sketched a generalization concerning certain classes of facts which have hitherto been treated disjointly. The basis of this unification is the category of head-transparency: heads are either transparent or non-transparent. Although this concept seems to be a coherent one, its definition remains rather informal. In the remaining time I would therefore like to discuss it in more detail. Before I begin, however, I should say that the clarifications I am going to propose will unfortunately not be entirely satisfactory either; in fact, in order to elucidate the concept of transparency I will have to take resort to a further intuitively coherent, nevertheless hard to define notion such as that of a quantifier. Despite this I hope that the discussion will not be completely trivial.

In a recent study which I have carried out together with Jean-Boase-Beier,[5] a generalization was proposed which says that suffixes typically establish a scope domain but are not able to assign a theta-role within that domain. It was therefore suggested that they best be characterized as a kind of operator. By and large, the proposal remains an intuitive one, yet at the logico-semantic level it leads to a natural contrast between quantifiers and non-quantifiers, i.e.,

[5] See Boase-Beier and Toman (1986)

between expressions that quantify and expressions quantified upon. However, at the level of syntactic representation, there is no one-to-one correlation between quantifiers and non-quantifiers on the one hand, and non-heads and heads on the other: both quantifiers and non-quantiers may function as syntactic heads, and both quantifiers and non-quantifiers may be specifiers.

A telling example for this comes from the syntax of NPs in Czech.[6] In this language, numerals equivalent to English cardinal numbers "one," "two," "three" and "four" are NP specifiers, and the objects counted are denoted by nouns in the head position. Hence we may say that the surface syntax of these phrases corresponds to what one may take to be the logical syntax of such phrases, i.e., the syntax of representation employed at the level of LF. Numerals from "five" to "ten", however, depart from this "normal form" in syntax in that they are arguably heads of syntactic phrases while the objects counted are genitives dependent on these numeral quantifiers. Incidentally, the latter type of numeral quantification has no special connotations, no special "partitive semantics." The only way to say "five men" in Czech simply is "five of men." To summarize the situation in syntax, we see that a quantifier of the same semantic type can either be realized as a non-head ("one" to "four") or as a head ("five" to "ten"). However, as far as the representation at the level of Logical Form is concerned, it can be argued that there are empirical reasons suggesting that numeral quantifiers always behave as non-heads at LF, whereas the nominal core of NPs in which they occur functions as the noun in the head position. To compensate this mismatch between LF and syntax, I have proposed a rule of re-analysis which, at the level of Logical Form, essentially supresses the head status of numeral quantifiers of the set "5—10" and yields something I will now call the Normal Form. That is, the proposal says that a syntactic form like (13) (a), in which "five" is the head, is re-analyzed, i.e., converted to the Normal Form (13) (b) at the level of Logical Form:

(13) (a) "(a) five of men"
(b) "five men"

One striking property of NPs with quantifier-like heads is that they are transparent with respect to extractions from NPs. For instance, a link between a clitic and its trace is well-formed in exampler of the following type:

(14) ... of-them$_{\text{clitic}}$ I saw [six t]$_{\text{NP}}$

[6] See Toman (1986) for details

The same link is however impermissible across a full nominal head:

(15) * . . . of-them$_{\text{clitic}}$ I saw [the end t]$_{\text{NP}}$

The suggestion that a Normal Form be constructed via a rule of re-analysis is a technical execution which seems necessary within the theory of binding. What is important in the present context, however, is that this kind of re-analysis only takes place if a transparent head is involved. It would thus seem that the notion of a transparent head is independently needed and that a transparent head is typically an item that defines a scope domain while not assigning a theta-role — in other words, a typical transparent head at the level of d-structure is an operator.

Pursuing this exposition further in an informal manner, I would also claim that the situation is in some sense analogous to the contrast between bridge verbs and their non-transparent counterparts. And, indeed, one might think of bridge-verbs as a kind of operator on propositions, i.e., attitude operators, whereas their concrete counterparts (such as *whisper*, etc.) denote an action (or a manner of action) and thus their semantic properties are not to be compared with those of bridge-verbs. Contrasts which I would claim are analogous to those between (4) and (6), for instance, can be constructed here too.

These final remarks should elucidate the relation between head-transparency and The Normal Form. It seems to me that whenever we compute the meaning of a complex expression with a transparent head, we basically suppress its head by relegating it to the operator status at the level of LF. In this sense, transparent heads always invite conversion to the Normal Form. Take German *bar* '-able, -ible'. The suffix is transparent, it maps the theta-grid of the verb in a "passive" manner, and has a modal meaning. By this token, *bar* is nothing like an operator with very simple properties, yet it is reasonable to any that *lesbar* means 'can be read'. If this is accepted then we see that *lesbar* does not exhibit The Normal Form in word-syntax since the head of this word is an operator at the level of LF. Hence a compensatory mechanism will be needed, be it a rule of re-analysis or some equivalent of it exploiting the transparency of the head.

The proposal which says that head-transparency should be correlated with operator status is no doubt an invitation to a very ambitious research program with many empirical problems, which I do not even dare to enumerate here, and some conceptual dangers as well. As for the latter, note that in order to make my approach really interesting, one would for instance like to know what a "(semantically) possible quantifier" is. As as far as I am familiar with the state of the art in this domain,[7] such notion is absent. Despite this, it seems to me that the proposal is a coherent one and worthy of further speculation.

[7] My thanks go to Peter Staudacher at this point

References

Boase-Beier, J. and Toman, J.: On Θ-Role Assignment in German Compounds. Folia Linguistica 20 (1986), 319–340.

Bergmann, R.: Verregnete Fereiengefahr und Deutsche Sprachwissenschaft: Zum Verhältnis von Substantivkompositum und Adjektivattribut. In: *Sprachwissenschaft* 5 (1980), 234–265.

Höhle, T.: Über Komposition und Derivation: Zur Konstituentenstruktur von Wortbildungsprodukten im Deutschen. In: *Zschr. Sprachwissenschaft* 1 (1982), 76–112.

Sandberg, B.: Der Bezug des Adjektivattributes bei substantivischen Zusammensetzungen. In: PBB (Tübingen) 106 (1984), 159–183.

Toman, J.: Wortsyntax. Eine Diskussion ausgewählter Probleme deutscher Wortbildung. Niemeyer, Tübingen 1983.

Toman, J.: Cliticization from NPs in Czech and Comparable Phenomena in French and Italian. In: Borer, H. (ed.): Syntax of Pronominal Clitics. Academic Press, New York 1986.

Williams, E.: On the Notions "Lexically Related" and "Head of a Word". *Linguistic Inquiry* 12 (1981), 245–274.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 221–234 (1986)

THETISCH vs. KATEGORISCH UND INFORMATIONSSTRUKTUR

MIORITA ULRICH
(Universität München, BRD)

0.1. Dt. *Es kommen die Tataren*, bzw. *Die Tatáren kommen/Die Tataren kommen bestimmt*; frz. *Il arrive un inspecteur (Il y a un inspecteur qui va venir)/L'inspecteur arrive aujourd'hui*; rum. *Arde o casă*, 'Es brennt ein Haus', 'Ein Háus brennt'/*Casa arde repede*, 'Das Haus brennt schnell (ab)'.

Die Äußerungen des ersten und die des zweiten Typs sind offensichtlich, auch schon für den naiven Sprecher, inhaltlich verschieden. Worin besteht aber der Unterschied?

0.2. In der bisherigen Literatur wurde meist angenommen, daß die Informationsstruktur der Äußerungen des ersten Typs (mit Verb-Subjekt-Anordnung) wie des zweiten (mit Subjekt-Verb-Anordnung) binär, zweigliedrig (Thema-Rhema, bzw. Topic-Comment) ist: der Unterschied würde also nicht die Informationsstruktur als solche, sondern nur die vom Subjekt und vom Verb in diesen Sätzen übernommenen Informationsrollen betreffen. Gemäß der gängigen Thema-Rhema-Theorie würde nämlich in einem einfachen Satz mit SV-Anordnung das Subjekt das Thema und das Verb das Rhema darstellen. In den Sätzen mit VS-Anordnung wäre hingegen das Verb das „Thema" und das Subjekt das Rhema. Für die Thema-Funktion des (intransitiven) Verbs werden meist zwei Arten von Kriterien aufgeführt: 1. semantische; 2. textkonstitutive. Einerseits sei die Nachstellung des Subjekts durch die semantische „Schwäche" des intransitiven („existentiellen") Verbs bedingt, das zu „vage" sei, um das Rhema bilden zu können. Andererseits sei die VS-Anordnung durch textkonstitutionelle Faktoren hervorgerufen: Die Textfunktion einer Äußerung mit VS-Struktur sei die „Präsentierung", die Einführung eines Text- oder Diskurs-Themas. Das semantisch „schwache" existentielle Prädikat hätte demnach auf der Satzebene die Funktion, ein rhematisches Subjekt einzuführen, das auf der Textebene das Thema des Diskurses darstellen soll.

0.3. Schwierigkeiten begegnet man bei dem Thema-Rhema-Modell vor allem bei der Interpretation der VS-Strukturen. Beim semantischen Kriterium kann man sich fragen, warum gerade das Verb als Thema interpretiert wird,

denn dies steht im Widerspruch zu der grundsätzlichen Annahme, daß Nominalglieder und nicht Verben ihrem Wesen nach dazu bestimmt sind, als Thema, d.h. als das, worüber etwas ausgesagt wird, zu fungieren. Ebenso kann das Ad-hoc-Kriterium der „semantischen Schwäche" des Verbs kaum als Begründung für die Thema-Funktion dieses Satzglieds angenommen werden, denn die nominalen Gruppen übernehmen die thematische Funktion keineswegs wegen einer für sie gar nicht erwogenen „semantischen Schwäche" (d.h. wegen einer quantitativen Eigenschaft), sondern vielmehr wegen ihrer Relationen in der Aussage. Darüber hinaus dürfen u.E. die sog. „semantische Schwäche" und die Aufgabe im Hinblick auf die Textkonstitution keineswegs als k o r r e l a t i v e Kriterien für dieselbe Struktur angenommen werden: Satz-Thema und Text-Thema stellen zwei verschiedene Größen dar, und außerdem kann ein Diskursthema ohne weiteres auch durch eine Äußerung mit tatsächlich binärer Thema-Rhema-Struktur, d.h. mit SV-Anordnung eingeführt werden. Mehr noch: die Entität, die in den Existentialkonstruktionen als „Rhema" eingeführt wird, braucht keineswegs auch das Thema des Diskurses zu bilden.

Ferner werden beim obigen Modell die intonatorischen Verhältnisse nicht berücksichtigt und daher auch nicht funktionell motiviert. Zwar weist eine Äußerung wie *Die Tatáren kommen* dieselbe SV-Anordnung wie *Die Tataren kommen (und. . .)* auf, sie zeichnet sich jedoch durch einen starken Satzakzent auf dem Subjekt aus und steht kennzeichnenderweise in Antwort auf die explizite (oder implizite) Frage *Was geschieht?* Bei einer genaueren Beobachtung fällt außerdem auf, daß in Antwort auf dieselbe Frage in anderen Sprachen wie Rumänisch, Italienisch, Spanisch, Russisch, Bulgarisch, Ungarisch, in der Regel eine Äußerung mit VS-Anordnung steht.

0.4. Auffassungen, wonach solche Äußerungen themalos und daher „gesamtrhematisch" seien, wird man ebensowenig beipflichten können, denn sie stellen diesen Typ von Äußerungen als defektive, nicht-vollständige „Varianten" der binären pragmatischen Gestaltungen dar. Sind sie es aber tatsächlich? Sie sind ja keineswegs den nur thematischen bzw. nur rhematischen Äußerungen analog, in denen das Thema bzw. das Rhema durch Kontext oder Situation gegeben ist und daher unausgedrückt bleibt.

0.5. Angesichts dieser und ähnlicher Schwierigkeiten will unser Beitrag zeigen:

A) Daß die sprachlichen Fakten, die mit der Informationsstruktur von Äußerungen zusammenhängen, nur dann zufriedenstellend analysiert werden können, wenn auch in diesem Bereich des Sprachlichen konsequent die von Eugenio Coseriu eingeführte Unterscheidung d r e i e r E b e n e n des Sprechens gemacht wird: des „Sprechens im allgemeinen", der „Einzelsprache" und des „Textes".

B) Daß bei einer so differenzierten Betrachtung sich eindeutig herausstellt:

a) daß es neben dem zweigliedrigen („kategorischen") Thema-Rhema-Typ noch einen anderen Typ mit eingliedriger („thetischer") Informationsstruktur gibt;

b) daß der Gegensatz zwischen beiden Typen als ein Universale des Denkens und des Sprechens angesehen werden muß, das in verschiedenen Sprachen durch verschiedene Verfahren und in verschiedenem Ausmaß zum Ausdruck kommt;

c) daß in vielen Sprachen der thetische Typ durch Verb-Subjekt-Anordnung ausgedrückt wird.[1]

1.1. Ausgangspunkt unseres thetisch-kategorischen Interpretationsmodells ist die Theorie des logischen Urteils von Franz Brentano und Anton Marty, die mit den dafür notwendigen Änderungen auf die sprachliche Äußerung schlechthin übertragen wird. In Anlehnung an F. Brentano — und im Gegensatz zur traditionellen Logik, die für alle Urteile eine zweigliedrige Struktur annahm — unterscheidet nämlich A. Marty (1918) zwei Haupttypen von Urteilen: A) ein zweigliedriges, teilbares und B) ein eingliedriges, nicht teilbares Urteil. Das zweigliedrige Urteil konstituiere sich durch die Verknüpfung zweier korrelativer Einheiten: eines (logischen) Subjekts und eines (logischen) Prädikats. Es enthalte demnach einen „subjektivischen" Teil, d.h. ein schon gesetztes Etwas („Subjekt"), von dem etwas prädiziert wird, und einen „prädikativischen" Teil („Prädikat"), d.h. etwas, was vom „Subjekt" prädiziert wird. Ein eingliedriges, einfaches Urteil bestehe hingegen einzig und allein im „Anerkennen" (bzw. „Verwerfen") eines „vorgestellten Inhalts", weshalb es auch weder ein logisches Subjekt noch ein logisches Prädikat enthalte. Das zweigliedrige Urteil nennt Marty k a t e g o r i s c h, d.h. „prädizierend", das einfache Urteil t h e t i s c h, d.h. „setzend", „feststellend", „(ein Faktum) behauptend". Dem kategorischen Urteil entspricht auf der sprachlichen Ebene eine k a t e g o r i s c h e A u s s a g e; so laut Marty: *Diese Blume ist gelb*, *Mein Bruder ist abgereist*, usw. Die sprachliche Entsprechung des thetischen Urteils ist die t h e t i s c h e A u s s a g e. Bei dieser unterscheidet Marty folgende Haupttypen: a) Impersonalsätze (z.B.: *Es regnet*, *Es sticht*, *Es klopft*, *Es brennt (in der Vorstadt)*, *Es schallt von den Zweigen*, *Es graut mir*, *Es sticht mich*, usw.); b) Existentialsätze (z.B.: *Es gibt gelbe Blumen*, *Es sind Menschen*, usw.); c) Universalurteile (z.B.: *Alle Dreiecke haben zur Winkelsumme zwei Rechte).*[2]

[1] Für eine ausführliche Behandlung der ganzen komplexen Problematik des Unterschieds thetisch/kategorisch und seiner verschiedenen Aspekte verweisen wir auf Ulrich 1985, wo auch eine Bibliographie zu diesem Thema zu finden ist.

[2] Wir selbst rechnen die Universalurteile ohne Ausnahme zu den "kategorischen", denn sprachlich sind sie, wie auch schon Kuroda 1973, 107f bemerkt hat, eindeutig zweigliedrig.

1.2. Martys Theorie des Urteils kann nun von der logischen Ebene des Urteils auf die Sprache übertragen werden, genauer auf die pragmatische Ebene der Mitteilung, d.h. der Darreichung von Information. Ähnlich wie bei Marty, wo wir als Oberbegriff das „Urteil" fanden, das erst in einem zweiten Schritt in ein zweigliedriges und ein eingliedriges eingeteilt wird, können wir in pragmatischer Hinsicht als Oberbegriff die „Äußerung" schlechthin als Einheit des Sagens nehmen, die dann in die zweigliedrige, d.h. thematisch-rhematische, und in die eingliedrige (d.h. nicht thematisch-rhematische und zugleich weder thematische noch rhematische) Äußerung unterteilt werden kann. Wir sprechen deshalb von „Äußerungen" und nicht von „Urteilen", von „Thema" und „Rhema" und nicht von „logischem Subjekt" und „logischem Prädikat". Wir unterscheiden also drei autonome Ebenen: a) des logischen Urteils (logisches Subjekt und logisches Prädikat), b) des einzelsprachlichen Satzes (grammatisches Subjekt und Prädikat), c) der Äußerung (zweigliedrige und eingliedrige informatorische Strukturen). So ist es für uns ohne weiteres möglich, daß z.B. eine Aussage als Satz grammatisches Subjekt, und grammatisches Prädikat aufweist, in pragmatisch-informatorischer Hinsicht aber eingliedrig ist (z.B. *Es regnet*), so wie andererseits ein grammatisch eingliedriger Satz pragmatisch gesehen zweigliedrig sein, d.h. eine Thema-Rhema-Gliederung aufweisen kann (z.B. rum. *vine*, '[er bzw. sie] kommt').

Wie tritt nun der Gegensatz thetisch-kategorisch, d.h. die Einteilung in eingliedrige und zweigliedrige Strukturen auf den Ebenen des „Sprechens im allgemeinen", der „Einzelsprache" und des „Textes"[3] auf?

2.1. Auf der Ebene des Sprechens (oder der Sprache) im allgemeinen, die grundsätzlich für alle Sprachen gilt und daher von jeder Einzelsprache absieht, tritt diese Einteilung als pragmatische („informatorische") Opposition auf. Sie betrifft die Bezeichnung, d.h. den Bezug auf die außersprachliche Wirklichkeit. Hier interessiert nicht, ob eine bestimmte Sprache diese Opposition auch wirklich sprachlich gestaltet und, falls ja, mit welchen sprachlichen Mitteln und in welchem Ausmaß.

Die Hauptopposition ist hier faktumbezogen/aktantenbezogen, d.h.: Eine thetische Äußerung setzt ein Faktum als Ganzes, in seiner Globalität, und weist aus diesem Grund eine eingliedrige informatorische Struktur auf, während eine kategorische Äußerung eine Aussage über einen Aktanten darstellt und deshalb zweigliedrig ist. Die korrelativen Glieder sind hier das Thema (= Aktant) und das Rhema (= Prädikation zu diesem Aktanten).

[3] Diese Unterscheidung übernehmen wir, wie w.o. erwähnt, von Eugenio Coseriu. Sie ist als ein Grundprinzip seiner Sprachtheorie anzusehen und liefert für ihn den methodologischen Rahmen grundsätzlich für alle Probleme der deskriptiven Sprachwissenschaft. Siehe insb. Coseriu 1975 und 1980.

2.2. Innerhalb des Thetischen können nun weitere Einteilungen festgestellt werden: das gesetzte „Faktum" kann das bloße D a s e i n (bzw. das „Nicht-Dasein") einer „Sache" oder aber ein E r e i g n i s sein. Diese Untertypen des „Faktumbezogenen" nennen wir d a s e i n s s e t z e n d (Existentialkonstruktionen) bzw. e r e i g n i s b e z o g e n (Äußerungen in Antwort auf die Frage „Was geschieht?"). Daseinssetzende und ereignisbezogene Äußerungen enthalten — außer den ereignisbezogenen Äußerungen ohne Agens, vom Typ *Es regnet* — selbsverständlich auch Aktanten; es wird jedoch nicht über sie „referiert". Vielmehr wird beim ersten Untertyp des „Dasein" des Aktanten gesetzt, während beim zweiten Untertyp der Aktant mit dem Ereignis eine unteilbare Einheit bildet: er ist sozusagen ein „Teil", ein „Aspekt" des Ereignisses. Bei dem ereignisbezogenen Untertyp kann allerdings eine weitere Einteilung vorgenommen werden: in einem Fall „wickelt" das Ereignis als solches den Aktanten ein und steht damit im Vordergrund der Mitteilung (z.B. *Es kommt ein Sturm*), in dem anderen Fall hingegen zieht der Aktant das Ereignis nach sich (sein Auftreten selbst stellt das Ereignis dar)[4] und er stellt sich somit in den Vordergrund (z.B. *Ein Stúrm kommt*!)

2.3. Auf der Ebene des Sprechens im allgemeinen sind also theoretisch folgende Möglichkeiten gegeben:

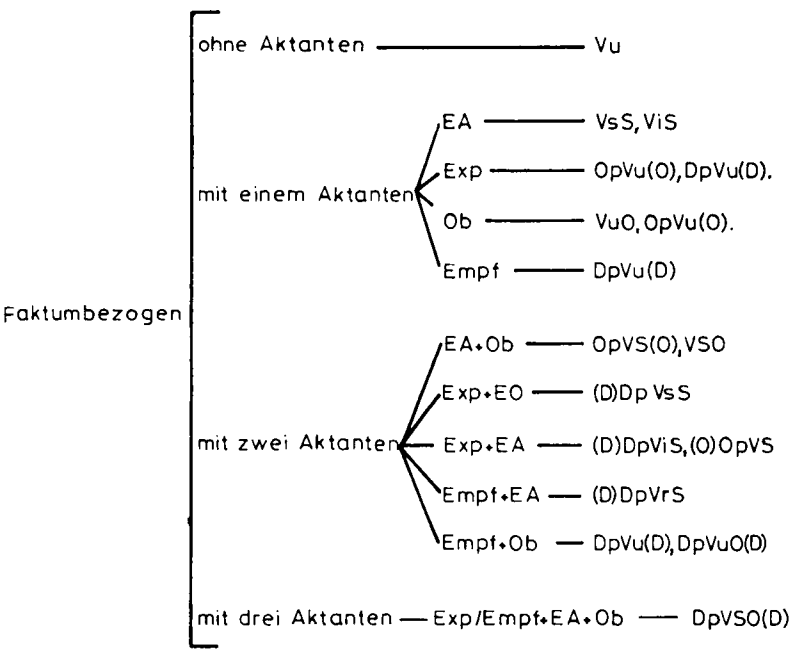


Abb. 1

[4] Deshalb kann in diesem Fall das grammatische Subjekt oft auch allein erscheinen: *Ein Sturm*! *Ein Schiff*! *Die Tataren*!

3.1. Die Unterscheidung t h e t i s c h / k a t e g o r i s c h würde allerdings nur die gedankliche, informatorische Struktur der Äußerungen betreffen und wäre dadurch nur für die Informationstheorie von Relevanz, wenn sie nicht in den historischen Sprachen mit einzelsprachspezifischen Ausdrucksmitteln als funktionelle Opposition gestaltet wäre. In den Sprachen könnte in der Tat diese Unterscheidung auch überhaupt nicht[5] oder nur zusammen mit anderen vorkommen. Dort, wo sie aber vorkommt — und dies ist b e i a l l e n von uns befragten Sprachen der Fall-, kann sie grundsätzlich in unterschiedlichem Umfang und mit unterschiedlicher einzelsprachlicher Strukturierung auftreten. Es gibt Sprachen, in denen — abgesehen freilich von den Sätzen ohne Aktanten, die immer thetisch, d.h. faktumbezogen sind —, die Unterscheidung nur in Sätzen mit einem einzigen Aktanten, d.h. nur in Sätzen mit Intransitivverben, gemacht wird. In anderen Sprachen kann die Unterscheidung auch bei Sätzen mit zwei oder sogar mit drei Aktanten vorkommen. So ist im Deutschen bei Aktivsätzen die Unterscheidung nur bei Sätzen mit dem Erstaktanten („Subjekt") wirklich üblich, weit weniger dagegen, bei Sätzen mit Subjekt und Objekt (s.w.u.). Im Französischen unterliegt die Unterscheidung in ihrer r e i n e n Form auch bei Sätzen mit einem einzigen Aktanten gewissen Restriktionen. Im Rumänischen hingegen ist die Unterscheidung auch bei Subjekt-Objekt-Sätzen vollkommen üblich; ebenso im Spanischen, wo sie — wie übrigens auch im Rumänischen —, sogar bei Sätzen mit drei Aktanten gemacht werden kann.

3.2. Überall aber ist in der Opposition thetisch-kategorisch das Kategorische das neutrale (extensive oder merkmallose) Glied; d.h. in Sätzen, wo die Unterscheidung in der betreffenden Sprache nicht gemacht wird und überall dort, wo sie nur fakultativ ist, tritt auch für thetische Bezeichnungen die kategorische Ausdrucksweise ein, nicht aber umgekehrt. Denn wie alle sprachlichen Oppositionen ist auch die Opposition thetisch-kategorisch eine inklusive, aufhebbare Opposition, die als solche in einem Text neutralisiert werden kann.

3.3. Andererseits können in den Sprachen innerhalb des Thetischen daseinssetzende und ereignisbezogene Äußerungen unterschieden werden, und bei den letzteren dazu noch diejenigen mit dem Ereignis bzw. mit dem Aktanten im Vordergrund. Oder die Unterscheidung kann den kategorischen Äußerungen nur faktumsetzende bzw. nur daseinssetzende Sätze als thetisch gegenüberstellen. So unterscheidet das Deutsche wenigstens einen Typ daseinssetzender Sätze vom allgemeineren Typ der ereignisbezogenen Sätzen und bei den letzte-

[5] Dieser Unterschied gehört nämlich zu den universellen M ö g l i c h k e i t e n der Sprache, nicht zu den wesentlichen, rational notwendigen Universalien (cf. Coseriu 1975b).

ren zwischen Ereignis im Vordergrund und Aktant im Vordergrund. Das Französische hingegen bietet das Faktumbezogene an erster Stelle als daseinssetzend und das Ereignisbezogene nur als einen nicht mehr „rein" thetischen Typ dar (s.w.u.). Im Rumänischen hingegen erscheint das Thetische an erster Stelle als faktumbezogen ohne Unterschied zwischen ereignisbezogen und daseinssetzend, d.h. die meisten daseinssetzende Sätze werden nicht anders als die ereignisbezogenen behandelt; die reinen daseinssetzenden Sätze stellen im Rumänischen einen marginalen Sondertyp dar, bei dem das Verb fast wie eine Interjektion funktioniert (s.w.u.). Wir wollen uns in dieser Hinsicht die Fakten in drei Sprachen, im Deutschen, im Rumänischen und im Französischen näher ansehen. Überall wird uns folgendes interessieren: (I) Die einzelsprachlichen Ausdrucksmittel, mit denen die Opposition realisiert wird, (II) Der Umfang der Opposition, d.h. die Zahl der einbezogenen Aktanten und (III) Die Einteilungen innerhalb des Thetischen.

4.1.1. Der Unterschied thetisch/kategorisch (bzw. faktumbezogen/aktantenbezogen) wird im Deutschen mit textgrammatischen Ausdrucksmitteln realisiert, und zwar einerseits und hauptsächlich durch die Anordnung der Satzkonstituenten (eine SV-(bzw. SVO-)Anordnung steht für eine binäre, thematisch-rhematische Informationsstruktur, während das VS-Muster auf eine eingliedrige, nicht teilbare Struktur hinweist),[6] andererseits aber mithilfe eines spezifischen Intonationsmusters, das bei einem bestimmten Untertyp des Thetischen vorkommt. Hinsichtlich des gut bekannten kategorischen, thematisch-rhematischen Typs, können, je nach Aktantenzahl und nach Thema-Wahl, die Anordnungen SV, SVO, OVS, SVDO, OVSD, DVSO (wobei D für das Dativ- und O für das Akkusativobjekt steht) erscheinen. Als Thema fungiert in allen dieser Äußerungen jeweils das erste nominale Satzglied, während der Rest der Äußerung rhematisch ist.

4.1.2. Wesentlich komplexer ist im Deutschen das System des Thetischen gestaltet, denn hier müssen zum einen (a) die Zahl der in thetischen Äußerungen einbezogenen Aktanten, zum anderen (b) die Untertypen des Thetischen berücksichtigt werden.

a) Regelmäßig wird der Unterschied thetisch/kategorisch nur bei Äußerungen ohne Aktanten *(Es regnet, Es wird getanzt)* und mit einem einzigen Aktanten gemacht. Letzteres kommt einerseits vor bei Konstruktionen mit intransitiven Verben *(Es kommen Kinder* bzw. *Kínder kommen* !), wo der einbezogene Aktant stets der Erstaktant ist, andererseits in unpersönlichen Kon-

[6] Die Konstruktionen mit inhaltsleerem nicht-kommutierbarem *es* (*es regnet*) und inhaltsleerem satzeröffnendem *es* (*Es kommen Kinder*) sind funktionell gleichwertig mit aktantenlosen bzw. VS-Konstruktionen.

struktionen mit Aktant (Typ *mich friert — es friert mich* bzw. *mir graust — es graust mir*), in denen der Erstaktant ein „Experiencer" (mit Akkusativ- oder Dativobjekt ausgedrückt) ist. Sätze mit transitiven Verben im Aktiv[7] sind in der Regel von der thetischen Konstruktion ausgeschlossen. Der Typ *Es sang ein Mädchen ein Lied* ist äußerst selten und klingt „gelehrt" und zugleich veraltet. Gegenüber der heutigen Umgangssprache sind in der Tat die entsprechenden Konstruktionen eindeutig als literarisch und archaisierend markiert; cf. folgende Belege aus der Rubrik „Unterm Strich" in der Illustrierten Stern (Nr. 21 vom 15.5.85, S.83):

Es hatte ein Maitanz in Denzen — *Es haben die Kinder in Glessen,*
Neun Monate drauf Konsequenzen — *Was Maikäfer sind, längst vergessen*

Ansonsten tritt im Deutschen in Konstruktionen mit mehr als einem Aktanten für thetische Bezeichnungen die kategorische Konstruktion mit SVO-Anordnung statt der zu erwartenden VSO-Anordung ein.

b) Bei Konstruktionen mit einem Aktanten zeichnet sich das Deutsche übrigens u.a. dadurch aus, daß es die „reine" Daseinssetzung *(es gibt* + Akkusativobjekt) von der Setzung eines Ereignisses eindeutig trennt. Außerdem macht das Deutsche, wie bereits erwähnt, den Unterschied „Ereignis im Vordergrund" und „Aktant im Vordergrund": *Es brennt ein Haus/Ein Haus brennt*![8] Beide Äußerungen stehen in Antwort auf die Frage *Was geschieht?* Die erste wird mittels VS- die zweite hingegen mittels SV-Anordnung ausgedrückt. Letztere weist freilich ein Intonationsmuster auf, das auch für eine andere Funktion stehen kann, und zwar für die kontrastive Hervorhebung des Themas in kategorischen Konstruktionen (Typ: *Das Háus brennt* (und nicht *die Scheune*). Die kontrastive Interpretation hängt allerdings stets vom weiteren Kontext ab, während dieselbe Äußerung kontextfrei nur als thetisch interpretiert werden kann.

4.1.3. Dies alles weist darauf hin, daß das Deutsche den Unterschied thetisch/kategorisch konsequent realisiert, wenn auch mit gewissen Ein-

[7] Zwei (oder sogar drei) autonome Aktanten können im Deutschen im Bereich des Thetischen in unpersönlichen Passivkonstruktionen erscheinen: *Es wurde dem Vater vom Sohn gedankt, Es wurden von Arbeitern viele Brücken gebaut,* usw. Die Passivkonstruktion scheint übrigens der Inbegriff des Thetischen zu sein, da sie das Ereignis vom Täter loslöst und es als "Faktum" darstellt.

[8] Auf eine ähnliche Idee, allerdings ohne Bezug auf Brentano und Marty ist bezeichnenderweise auch A. W. de Groot: Les oppositions dans les systèmes de la syntaxe et des cas". In: Mélanges de linguistique, offerts à Charles Bally. Genève 1939, 115, 117 gekommen, der Ausdrücke wie *Das Haus brennt!* als "semantisch" eingliedrig und nicht prädikativ interpretiert. Er ist der Meinung, daß die Satzbetonung in solchen Fällen die prädikative Funktion des Verbs aufhebt und daß das Verb selbst dem "sujet sémantique" einverleibt wird.

schränkungen, was die Zahl der in thetischen Äußerungen einbezogenen Aktanten betrifft. Außerdem kennt es fast alle Untertypen des Thetischen.

4.2.1. Das Rumänische unterscheidet sich in diesem Bereich vom Deutschen hauptsächlich dadurch, daß es die Grenzen der Opposition wie folgt zieht: a) es macht den Unterschied regelmäßiger und im höchsten Ausmaß, was die Zahl der in thetischen Konstruktionen zugelassenen Aktanten betrifft; b) es kennt jedoch nicht den Unterschied zwischen „daseinssetzend" und „ereignisbezogen"; c) die einzelsprachlichen Ausdrucksmittel sind ausschließlich positioneller Natur (die SV- bzw. SVO-, OVS- etc. Anordnung steht für das Kategorische, die „invertierte" Reihenfolge VS,[9] VSO, VSDO für das Thetische. Das Rumänische kennt neben den unpersönlichen Konstruktionen ohne Aktanten (*ninge* „es schneit") und den persönlichen und unpersönlichen mit einem Aktanten, auch thetische Konstruktionen mit zwei und sogar mit drei Aktanten. In den thetischen Konstruktionen mit einem einzigen Aktanten kann dieser der Erstaktant (*Vine un vapor* 'es kommt ein Schiff'), der „Experiencer" (*Mă doare* 'Ich habe Schmerzen'; *Mi-e foame* 'Ich habe Hunger'), der „Patiens" (*Îl plouă* [*în cap*] 'Es regnet ihm [auf den Kopf]' oder der „Empfänger" (*Îi plouă omului*[*în casă*] 'Es regnet dem Mann [ins Haus]' sein. In den Konstruktionen mit zwei Aktanten treten folgende Paare auf: Erstaktant und „Patiens" (*Ceartă mama băiatul* 'Die Mutter tadelt den Buben'), „Experiencer" und „Erfahrenes Objekt" (*Mi-e o foame grozavă* 'Ich habe einen schrecklichen Hunger'), „Experiencer" und Erstaktant (*Mă doare capul* „Mir tut der Kopf weh") „Empfänger" und „Objekt" (*I se dăruiesc mamei flori* 'Der Mutter werden Blumen geschenkt'), „Emfänger" und Erstaktant (*I s-a predat generalul X* 'Der General X hat sich ihm ergeben'). Die thetischen Konstruktionen mit drei Aktanten (*I-a făcut mama prăjituri băiatului* 'Die Mutter hat dem Jungen Kuchen gebacken') sind zwar selten, jedoch keineswegs ungrammatisch. Sie werden allerdings oft durch kategorische Konstruktionen ersetzt (SVOD).

4.2.2. Das Rumänische sieht von jeder Einteilung des „Faktumbezogenen" in die Bezeichnungstypen „daseinssetzend"[10] und „ereignisbezogen" ab und drückt beides durch VS-Anordnung aus. Daraus ergibt sich, daß es ebenso-

[9] Die VS-Anordnung steht übrigens im Rumänischen — allerdings ausschließlich wenn mit bestimmten intonatorischen Mitteln gekoppelt — auch für den Ausdruck der Abtönung. Näheres hierzu in Ulrich 1985, 256–284 und Ulrich 1985b.

[10] Für die "reine" Daseinssetzung kennt das Rumänische die transitiven Kontruktionen mit den "präsentativen" Elementen *iată* 'da', *uite* 'sieh mal', *na* 'da [hast du ihn]'. Es handelt sich hierbei um Interjektionen, die hier jedoch wie transitive Halbverben funktionieren und als solche auch eine Verbalrektion aufweisen (z. B. *iată-l* 'Da ist er', wörtl. 'Da ihn'; cf. it. *Ecco Giovanni* !, span. *He aquí a Juan* !, frz. *Le voilà* !). Diese Konstruktionen sind immer thetisch, denn sie setzen das Dasein einer "Sache" ohne darüber etwas zu prädizieren.

wenig den Unterschied zwischen „Ereignis im Vordergrund" und „Aktant im Vordergrund" macht, so daß rum. *Vin cazacii* sowohl dt. 'Es kommen die Kosaken', als auch 'Die Kosáken kommen' entsprechen kann. Dies alles zeigt, daßdie Einteilung im Rumänischen nach der Zahl und der Art der einbezogenen Aktanten erfolgt, wie auch der folgenden schematischen Darstellung entnommen werden kann (cf. Ulrich 1985, 74):
(Wir bezeichnen mit Vs das Verb *sein*, mit V das transitive, mit Vi das intransitive, mit Vu das unpersönliche und mit Vr das reflexive Verb; EA bezeichnet den „Erstaktanten", d.h. den ersten („realen") Bezug des Verbs *sein* und der intransitiven und reflexiven Verben und den „Agens" in der aktiven Konstruktion der transitiven Verben; Ob steht für das „Objekt" oder den „Patiens" einer Handlung, Exp für den „Experiencer" und Empf für den „Empfänger"; EO ist das von einem „Experiencer" erfahrene „Objekt", D das Dativ- und O das Akkusativobjekt, Dp das pronominale Dativ- und Op das pronominale Akkusativobjekt. In Klammern stehen die grammatischen Aktanten, die erscheinen k ö n n e n, ohne Klammern diejenigen, die erscheinen m ü s s e n; so können z.B. der „Experiencer" und der „Empfänger" zwar auch nominal ausgedrückt werden, aber auch in diesem Fall müssen sie auch pronominal „angekündigt" bzw. wiederaufgenommen werden; cf. *îl doare — îl doare pe Ion*).

4.3.1. Das Französische macht den Unterschied thetisch/kategorissch einerseits durch die Reihenfolge der Satzglieder, andererseits — und diesmal im Unterschied zum Deutschen und Rumänischen — durch gewisse Konstruktionen, die bei einiger Untertypen des Thetischen vorkommen.

Der Verwendungsbereich des Thetischen ist im Französischen weit weniger umfangreich als im Rumänischen: es gibt thetische Äußerungen ohne Aktanten *(Il pleut)*, Äußerungen mit einem einzigen Aktanten (daseinssetzende und ereignisbezogene Äußerungen mit existentiellen und intransitiven Verben) und einen einzigen Typ von thetischen Äußerungen mit zwei Aktanten. Bei den daseinssetzenden Konstruktionen können zwei Typen unterschieden werden: a) Existentialkonstruktionen im engeren Sinne, mit dem Verb *sein (il est un X, il y a X, il existe X)* und b) Konstruktionen mit einer begrenzten Anzahl von intransitiven und „reflexiven" Verben *(venir, arriver, se produire, se présenter, se préparer* [*Il se prépare un orage*]). Bei den ereignisbezogenen Äußerungen mit einem Aktanten kennt das Französische nur den Typ mit dem Aktanten im Vordergrund (s. dt. *Ein Stúrm kommt !)*, da die Konstruktion mit unpersönlichem Verb (**Il crie un enfant*) in diesem Fall ausgeschloßen ist und die Konstruktion mit SV-Anordnung (*Un enfant crie*) nur kategorisch sein kann (die zulässigen Konstruktionen mit unpersönlichen Verben — *Il vient un orage, Il souffle un vent terrible*, etc. — gehören u.E. zum daseinsetzenden Typ). Bei den ereignisbezogenen Äußerungen mit dem Aktanten im Vorder-

grund handelt es sich um Konstruktionen vom Typ *C'est la maison qui brûle* ![11] 'Das Háus brennt!', freilich nur wenn diese Äußerung in Antwort auf eine Frage vom Typ *Qu'est-ce qui se passe?* steht oder aber kontextfrei ist. Die Bestandteile dieser Konstruktion sind: a) die Existentialkonstruktion mit *c'est*, die die Funktion hat einen Aktanten (meistens das Subjekt des Satzes) zu „setzen"; und b) ein an diesen Aktanten angeschlossener Relativsatz. Dieser Aktant wird allerdings nicht allein (wie dies im Falle der kontrastiven und kategorischen Funktion der Konstruktionen mit *c'est. . .qui* erfolgt!), sondern immer als Bestandteil eines Faktums „vorgestellt", denn hier wird ein ganzes Faktum „gesetzt", d.h.:

nicht:	*C'est*	*maman*	*qui arrive*	'Die Mutter kommt [nicht jemand anders]'
sondern:	*C'est:*	*maman qui arrive*		'Es kommt die Mutter (Die Mútter kommt!)'

Die Gesamtkonstruktion vom Typ 2 ist demnach als thetisch zu interpretieren;[12] sie enthält eine kategorische Komponente (den Relativsatz):

THETISCH	
C'est maman	*qui arrive*
KATEGORISCH	

Dies bedeutet, daß im Französischen das Ereignis eigentlich als eine Daseinsetzung präsentiert wird und zwar mit dem Ereignis (*la maison qui brûle, maman qui arrive*) als Relativsatz.[13]

4.3.2. Nicht anders verhält es sich, was die Informationsstruktur betrifft, mit den thetischen Äußerungen mit zwei Aktanten, denn auch hier ist die Gesamtkonstruktion als thetisch zu interpretieren, sie enthält aber eine kategorische Komponente in der Form eines eingebetteten Relativsatzes. Solche Konstruktionen mit zwei Aktanten („Experiencer" und „Objekt") liegen im Französischen in Äußerungen vom Typ *J'ai mon mari qui est malade* vor. Diese Periphrasen mit *avoir* stehen entweder in Antwort auf die Frage *Was*

[11] Außer der Konstruktion mit *c'est . . . qui* gibt es im Französischen noch weitere Konstruktionen, die ein Faktum präsentieren können, nämlich: *Il y a X qui, Voilà X qui, Et X qui* (cf. Ulrich 1985, 97—110).

[12] Übrigens kann nur dieser zweite Typ auch ohne *c'est* erscheinen (*Le rôti qui brûle! La branche qui casse!*), nie aber der erste, wo *c'est* als Zeichen der Konstrastivität obligatorisch ist.

[13] Es muß allerdings auch darauf hingewiesen werden, daß das Französische für derartige pragmatische Inhalte auch kategorische Ausdrücke (mit SV-Anordnung) verwendet, d. h., daß die entsprechende Opposition in einzelsprachlicher Hinsicht oft aufgehoben wird.

geschieht? oder als Begründung, als Erklärung für ein bestimmtes Verhalten; z.B. [*Je ne peux pas venir,*] *j'ai la tête qui* (*me*) *tourne, j'ai mon frère qui vient d'arriver,* etc. Die Verwendung dieser Periphrasen ist allerdings auf diejenigen Fälle beschränkt, in denen der zweite Aktant (= das grammatische Objekt) (*la tête, mon frère*) mit dem „Experiencer" (= Subjekt von *avoir*) in einer sog. „relation inaliénable" steht. Dieser Zweitaktant fällt nämlich außersprachlich z.T. mit dem „Experiencer" zusammen (z.B. Körperteile), oder er gehört eng zu seiner „vitalen" Sphäre (Eltern, Geschwister, Kinder, nähere Verwandte). Liegt keiner dieser Fälle vor, so wird die Konstruktion mit *j'ai. . . qui* durch *il y a. . .qui* ersetzt (z.B. *Il y a un monsieur qui vient d'arriver*), wodurch die Konstruktion allerdings um einen Aktanten reduziert wird. Auch werden solche „thetischen" Periphrasen in der gehobenen Sprache oft durch kategorische Konstruktionen ersetzt, z.B. *Mon frère vient d'arriver,* anstelle des umgangssprachlichen und familiären *J'ai mon frère qui vient d'arriver.*

4.4. Unter den romanischen Sprachen weichen in diesem Bereich das Französische und das Rumänische am weitesten voneinander ab. Das Italienische (das Gemeinitalienische) und das Portugiesische befinden sich etwa auf dem halben Weg zwischen diesen beiden Sprachen. Sie unterscheiden zwar streng zwischen thetisch und kategorisch bei den Konstruktionen mit einem einzigen Aktanten (it. *Arriva una nave/La nave arriva oggi* bzw. port. *Chega um barco/O barco chega hoje*), bei Äußerungen mit zwei Aktanten (Subjekt und Objekt) ziehen sie aber die kategorische Konstruktion vor. Fast genau wie das Rumänische und nur mit geringen Unterschieden im einzelnen verhält sich hingegen das Spanische,[14] z.B.:

Spanisch	Rumänisch
llueve	*plouă*
llega un barco	*vine o corabie*
murió Pedro	*a murit Petru*
se abre la puerta	*se deschide ușa*
cae la noche	*se lasă noaptea*
me duele la cabeza	*mă doare capul*
me gusta la muchacha	*îmi place fata*
se nos heló el corazón	*ne-a înghetat inima*
la ha mordido un perro a mi madre	*a muscat-o un câine pe mama*
ha comido el perro al gorrión	*l-a mîncat câinele pe vrăbiete*
le ha dado la madre una bofetada al niño	*i-a dat mama o palmă băiatului*

[14] Wir haben in unserer o.a. Dissertation (S. 218–223) diese auffallende Übereinstimmung insbesondere auf das Vorhandensein des markierten persönlichen Objekts und der Objektkonjugation in beiden Sprachen zurückgeführt.

5.1. Auf der Ebene des „Textes" ist vor allem die Art hervorzuheben, wie die thetischen und kategorischen Konstruktionen zur K o n s t i t u t i o n von umfassenderen Texten („Makrotexten") beitragen (auf andere keineswegs unwichtige Textfunktionen können wir hier nicht eingehen).[15] Als Textopposition läßt sich in dieser Hinsicht insbesondere die Opposition p r ä s e n t a t i v / k o n t i n u a t i v feststellen. Als präsentativ (einführend) funktionieren die eingliedrigen, thetischen Äußerungen, da sie ein Faktum als solches „setzen"; als kontinuativ (d.h. als wiederaufnehmend), weiterführend hingegen die kategorischen Aussagen, da sie etwas zu einem (meist) schon gegebenem Thema aussagen. Während die thetischen Äußerungen ohne Aktanten den Rahmen und die Umstände für kategorische Äußerungen bilden (z.B. *Es hagelte fürchterlich. Peter ging trotzdem hinaus. . .*), „präsentieren" die faktumbezogenen (daseinssetzenden und ereignisbezogenen) Äußerungen mit Aktanten ein Faktum und z u g l e i c h ein (eventuelles) Thema eines ganzes Makrotextes oder eines Textabschnittes („Diskursthema"). Z.B.: *Es war einmal ein König* [thetisch]; *Der König hatte drei Töchter* [kategorisch]. *Der Húnd ist weggelaufen* [thetisch]; *Er war nicht richtig angebunden* [kategorisch].

5.2. Im Rahmen bestimmter Texttraditionen fangen gewisse Texte tatsächlich mit einer präsentativen Formel (vom Typ etwa: *Es war einmal. . .*) an; in solchen Fällen beginnt die erzählte Geschichte (das, was die russischen Formalisten *Sujet* nennen) v o r der „realen", zu erzählenden Geschichte (bei den russischen Formalisten: *Fabula*).

Die erzählte Geschichte kann aber auch gleichzeitig mit der Fabula anfangen (*Ein König hatte eine Tochter*). In diesem Fall fängt der Text mit einer kategorischen, „kontinuativen" Konstruktion an, die gleichwohl ein Diskursthema einführen kann (der *König* oder seine *Tochter*). Auch in diesem Fall ist die Opposition präsentativ/kontinuativ eine „inklusive"; das Kategorische (kontinuative) funktioniert eben auch hier als „extensives" Glied.

Die erzählte Geschichte kann allerdings auch mit einer kategorischen Konstruktion wie der folgenden anfangen: *Den Arm bitte langsamer bewegen.* In diesem Fall wird man — offensichtlich absichtlich — gleich „in medias res" geführt, denn hier fängt die erzählte Geschichte n a c h dem Beginn der „realen" Fabula an.[16]

[15] Thetische und kategorische Äußerungen können freilich auch allein ganze Texte („Mikrotexte") darstellen und somit autonome „Sinneinheiten" ausdrücken (Feststellung, Erwiderung, Einwand, Drohung, Ironie, Anspielung, Unterstellung, etc.).

[16] In bestimmten Textsorten wird das Diskursthema oft durch den Titel eingeführt (z. B. in einem Buch über Igel), und der ganze Text wird nur noch kategorisch konstruiert; und in der Ich-Erzählung erübrigt sich die Einführung des Diskursthemas (d.h. des Protagonisten) von vornherein.

Literaturhinweise

Coseriu, E.: Determinierung und Umfeld. Zwei Probleme einer Linguistik des Sprechens. In: Coseriu, E.: Sprachtheorie und allgemeine Sprachwissenschaft. München 1975, 253–290.

Coseriu, E.: Die sprachlichen (und die anderen) Universalien. In: Schlieben-Lange, B. (Hrsg.): Sprachtheorie. Hamburg 1975, 127–161.

Coseriu, E.: Textilinguistik. Eine Einführung. In: Tübinger Beiträge zur Linguistik 109 (1980).

Kuroda, S.-Y.: Le jugement catégorique et le jugement thétique. Exemples tirés de la syntaxe japonaise. In: Langages 8 (1973), 81–110.

Marty, A.: Über subjektlose Sätze und das Verhältnis der Grammatik zu Logik und Psychologie. In: Marty, A.: Gesammelte Schriften. Vol II. Halle a.d.S. 1918, 3–309.

Marty, A.: Über die Scheidung von grammatischem, logischem und psychologischem Subjekt, resp. Prädikat. In: Marty, A.: Gesammelte Schriften. Vol. II. Halle a. d.S. 1918, 309–364.

Ulrich, M.: Thetisch und Kategorisch. In: Romanica Monacensia 24 (1985).

Ulrich, M.: La tonalisation syntaxique en roumain. In: International Journal of Rumanian Studies 4 (1984–1986) : 2. Amsterdam 1985a. 81–88.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 235–244 (1986)

THE ORGANIZATION OF THE LEXICAL COMPONENT: NOUN-COMPOUNDS IN SPANISH

SOLEDAD VARELA

(University of Madrid, Spain)

1. Structural analysis

The structure of Spanish Noun-Compounds, traditionally described as having the form [V + N], like *paraguas* in (1), poses a number of problems:

$$[+\text{ plur}]_N$$

(1) $[[\text{para}]_V\ [\text{aguas}]_N]_N$ = 'umbrella' (lit. 'stop waters').
$[+\text{ plur}]$

a) The place of the lexical head; that is, whether it is inside the compound (endocentric compound) or outside (exocentric compound).

b) The category feature of the first constituent; whether it is [+ V] or [+ N].

c) The inflectional markers realized inside the compound, and markers realized outside it and attributed to the complete word.

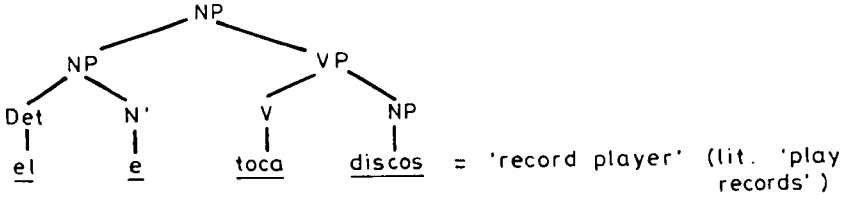
d) Diminutive affixation attached to the second constituent but consistently referring to the compound as a whole.

The same problems appear in other Romance languages such as Italian, French or Catalan.

Recent research on these constructions has mainly focused on items (a) and (b), although the questions raised in (c) are always considered.

In Contreras (1985), Spanish verb + noun compounds (2) are analyzed as having a head, an empty category located outside the compound:

(2) Sp.



(Contreras 1985)

Another exocentric interpretation, referring this time to Italian compounds, is found in Scalise (1984). According to his analysis, reproduced in (3), the head of the compound — not explicitly defined as an empty category — appears to the right of the compound, in accordance with the "Righthand Head Rule" (RHR) of Williams (1981).

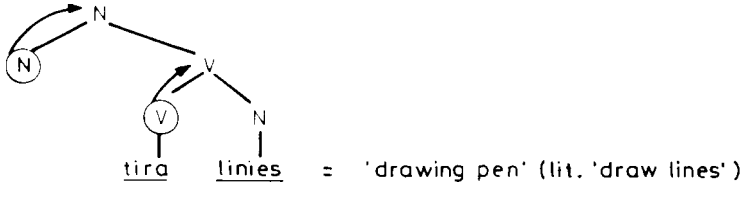
(3) It.

$$\underset{N}{[}[\text{porta}]_V\ \underset{N}{[\text{cener}]}]\underset{N}{+i}\underset{[+\text{plur}]}{]} = \text{'ashtrays' (lit. 'carry ashes')}$$

(Scalise 1983, 1984)

Finally, the analysis by Cabré and Rigau (1986) of the Catalan compounds in (4) is another instance of an exocentric interpretation. The head appears to the left of the compound, as an empty node.

(4) Cat.



(Cabré & Rigau 1986)

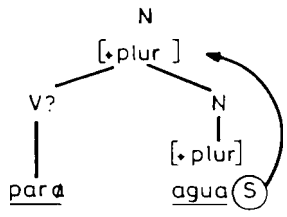
I will not deal here with the question of the exocentricity or endocentricity of these compounds, nor will I treat the question of the syntactic category of the first constituent, a matter directly connected with the first issue. I am going to focus on the question of inflectional and evaluative affixation in relation to these nominal compounds in Spanish, since these remain problematic questions in the analyses referred to above.

2. Inflectional affixation and nominal compounds. Morphological vs. syntactic inflection

Most of these compounds have a second constituent [+ N] with the plural marker *-s*; this noun is generally interpreted as a "complement" of the first constituent, as (2) shows.

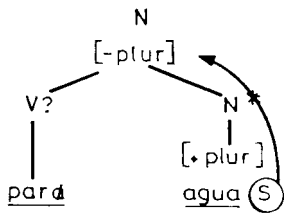
The inflectional affix, while always occurring in a peripheral position, can be interpreted, though, either as syntactically relevant (5a) or not (5b).

(5) (a)



'umbrellas'

(5) (b)



'umbrella'

This is the reason why, as shown in (6), we can get the compound *paraguas* to agree either with a Det [+ plur]: *los paraguas* or with a Det [— plur]: *el paraguas*, without any change:

(6) [el paraguas]$_{NP}$ [los paraguas]$_{NP}$
[—plur] [+ plur]
'the umbrella' 'the umbrellas'

The possibility of assigning the inflectional affix either to an internal constituent or to the whole compound is not rare in a level order morphology, where inflectional markers have been recognized as having syntactic significance: (5a), while others are interpreted as instances of morphological inflection: (5b) (see Hammond 1984).

I will now present some data from Spanish to support this interpretation:

a) There are certain formations of the type I am analyzing in which the noun-complement appears in the singular. Such nouns are perceived as archaic or dialectal forms and, if preserved in the active vocabulary, normally acquire the plural marker *-s*; I give some examples in (7).

(7) *buscapleito* → *buscapleitos*: 'troublemaker'
(old) (new)
abreboca (dial.) <Ecuador & Venezuela>: 'appetizer'
ganapán (lost): 'messenger'
(Lloyd 1968).

b) In other cases, both forms are attested, but with some semantic difference, so that the "blocking principle" defined in Kiparsky (1982) need not apply. The plural form is more recent and very often semantically transparent; that is, its meaning can be obtained compositionally from the meaning of its constituents. The compounds with a singular second element exhibit a greater degree of metaphoric deviation, close to total lexicalization; (see (8)).

(8) *tapaboca*: 'objection, refutation, argument'.
tapabocas: '[garment] that covers (*tapa*) the mouth (*bocas*)', 'scarf'
buscavida: 'diligent, quick, astute'
buscavidas: '[person] who pries into (*busca*) someone's affairs' (lit. 'lives' = *vidas*)'.

c) Moreover, the presence or absence of the plural marker allows us to distinguish the coordinate interpretation from the subordinate one in which the second element acts as a complement of the first. In (9) we have *cazabombardero* = 'fighter-bomber', a coordinate compound with the second noun in the singular, that becomes a subordinate compound: *cazabombarderos* = 'plane that tracks down bombers' when the second constituent is pluralized, after he model of *cazasubmarinos* = 'boat that tracks down submarines'

(9) *cazabombardero* = 'fighter-bomber'
$[[\text{caza}]_N + [\text{bombardero}]_N]_N$
cazabombarderos = 'plane that tracks down (lit. 'hunts' = *caza*) bombers (*bombarderos*)'
cazasubmarinos = '[boat] that tracks down submarines'

d) The second noun appears invariably in the singular only as a consequence of its having the feature [— count] (see 10):

(10) *limpianieve* = 'snowplough' (but cfr. *quitanieves*)
guardapolvo = 'dust coat'
pasatiempo = 'pastime'

So, when plural *-s* appears in such cases, the only possible agreement would be with a Det [+ plur]: *los*; that is, the plural marker *-s* is always external in those compounds, and therefore syntactically relevant.

Summarizing, the data just presented show lack of plural in the second element, as well as a difference in semantic interpretation and in the syntactic relation between the constituents, both governed by the presence or absence of the plural affix. In other words, there is evidence in Spanish to justify the distinction proposed by Hammond (1984) between morphological and syntactic inflection.

3. Evaluative affixation in the nominal compound

The evaluative affixation often falls on the borderline between derivation and inflection, as for instance, in Scalise's (1985) model. On the one hand, it resembles inflection in not changing the syntactic category of the base form;

on the other, it resembles derivation because it always appears preceding the inflectional affixes, and can sometimes even go inside other derivational affixes.

Spanish diminutive *-it-* is generally considered an infix (see, more recently, Jaeggli, 1980) as the examples in (11) show:

(11) *azúcar* $\overset{\text{dim.}}{\longrightarrow}$ *azuqu -it- ar* = 'little sugar'

Carlos $\overset{\text{dim.}}{\longrightarrow}$ *Carl -it- os* = 'Charlie'

anarqu⁻ista]$_{DS}$ $\overset{\text{dim.}}{\longrightarrow}$ *anarquist -it- a* = 'little, insignificant anarchist'

As has been traditionally observed, the Spanish diminutive has the ability to change the gender marker of the base, "recovering" so to speak the prototypical or canonical gender allomorphs of the masculine (*-o*) and the femenine (*-a*), whenever such forms are not manifested in the base; see for examples (12):

(12) *man*ⓞ$_N$ $\overset{\text{dim.}}{\longrightarrow}$ *man-it-*ⓐ$_N$ = 'little hand'
[— masc] [— masc]

jefⓔ$_N$ $\overset{\text{dim.}}{\longrightarrow}$ *jef-ec-it-*ⓞ$_N$ = 'little chief'
(+ masc] [+ masc]

*señal*ø$_N$ $\overset{\text{dim.}}{\longrightarrow}$ *señal-it-*ⓐ$_N$ = 'little signal'
[— masc] [— masc]

*canal*ø$_N$ $\overset{\text{dim.}}{\longrightarrow}$ *canal-it-*ⓞ$_N$ = 'little channel'
[+ masc] [+ masc]

In compound nouns, the diminutive *-it-* behaves also as an infix applied to the entire compound, as attested by the semantic interpretation of the nouns in (13):

(13) *tocadisquitos* = 'little recordplayer' (*player of little records)
paragüitas = 'little umbrella' (*object that stops little waters)

In such cases, we can, however, observe several different anomalies:

a) the diminutive fails to affect the gender marker of the compound — always masculine — as might be expected from the behavior of the non-

complex words in (12). That is, we never get nouns like the ones in (14), possible words since — as I have said — the entire compound is always masculine.

(14) **paragü-it-os, *cuentagot-it-os, *apagavel-it-os, . . .*
'umbrella' 'dropper' 'candle extinguisher' (+ little)

Moreover, in the case of the compounds, the allomorphic rules that apply to the rest of the diminutive formations are not observed. As can be seen from the second word of (12) — leaving aside other stipulations- when a word is disyllabic and ends in *-e*, the diminutive infix *-it-* is supplemented in Spanish with another suffix (or "interfix") *-ec-*. This rule applies also to monosyllables but never to trisyllabic words; see examples in (15):

(15)

choque —dim.→	*choqu-ec-it-o*	'impact'	(+ little)
sol —dim.→	*sol-ec-it-o*	'sun'	
carrete —dim.→	*carret-it-o*	'reel'	

Interestingly, the diminutive compound noun in (16), while having more than two syllables, still takes the *-ec-* augment. This happens in all compounds where the second element is one of the nouns that require *-ec-* owing to their syllable structure:

(16) *parachoque* —dim.→ *parachoqu-ec-it-o* (**parachoqu-it-o*)
'fender'

The gender marker anomaly, as well as the diminutive allomorphic irregularity, both show that the evaluative affixation is realized over the second constituent of the compound, which is quite inconsistent with the semantic interpretation assigned to it — see again the glosses in (13). We have here, then, a case of discrepancy between formal structure and semantic interpretation.

b) There is also a relevant anomaly with respect to the inflectional affix of number, in those cases where the *-s* marker is attributed to the second constituent, i.e., in the cases considered instances of "inflectional morphology". This anomaly arises from the fact that the diminutive, while referring to the entire compound, appears before the plural suffix of a constituent placed in the innermost layer of structure: *agü-it-as*.

That is, we would expect forms like the ones in (17) which, in fact, do not exist. Here again is another case of a "paradox" between formal structure and semantic interpretation:

(17) **paraguas-it-o(s)*
**cuentagotas-it-o(s)*

Such incongruities with regard to the ordering of evaluative affixation pose a further problem for two well-known restrictions on the attachment of affixes, namely the "Atom Condition" (Williams 1981) and the "Adjacency Condition" (Siegel 1977), independently of the place of the head of the compound, be it inside or outside the word, and either to the right or to the left of it.

In the case of the "Adjacency Condition", this is because, as it is obvious, adjacency is lost because of the interference of some inflectional suffixes which refer to an internal constituent. In the case of the "Atom Condition", the problem is due to the following questions:

a) none of the analyses proposed for those compounds takes the second constituent as the head or nucleus of the word, since its inherent features very often do not coincide with the features of the compound noun. In the case of the singular noun: *el paragüitas*, we would like the evaluative affix, realized on the second element, to percolate to the top of the word, but not the inflectional affixes of gender and number attached after it.

b) in the case of the plural noun, *los paragüitas*, we would also like the number suffix to percolate, with the evaluative affix, but not the gender marker.

No matter which element is the head of these compounds, the "atom condition" is severely weakened if some internal affixes are allowed to percolate to the top of the word, while others, realized on more external layers, are not.

As can be appreciated, the data presented thus far pose a number of problems for a level order morphology like that defended, for instance, in Scalise (1985).

4. Morphological and phonological derivation

Evaluative affixation seems to be a case of derivation. It is not comparable to inflection from a paradigmatic point of view: there are no "evaluative classes" in the manner of "inflectional classes".

From a syntagmatic point of view it cannot be treated like inflection either: evaluative affixes are not dependent on syntactic structure or grammatical relations, as inflectional affixes normally are. In Anderson's terms, evaluative affixes -at least in the Romance languages- are not relevant to

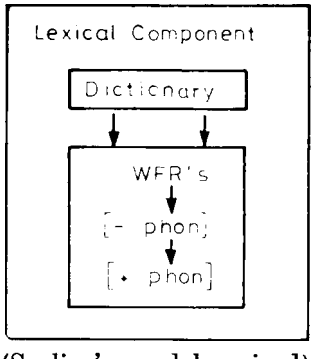
syntax, hence they are not part of inflection. Moreover, evaluative affixation does affect semantic interpretation in some way, whereas inflection does not.

Nevertheless, as I have said, in some respects they behave very similarly. In my opinion, the similarity is more "phonological" than "morphological" in nature. Both types of affixation are linked to certain rules of the phonology proper, which do not apply to the rest of the morphology and both are immune to certain structural conditions of a purely morphological character (see Anderson 1982, and his proposal for including the rules of inflection inside the phonological component).

Thus, while I consider that in a model like Scalise's (1984), evaluative affixation should be classified inside derivation, I suggest distinguishing between phonological and non-phonological derivation, a distinction which is complemented by that between morphological and syntactic inflection mentioned above. The WFR subcomponent inside the lexical component will then be organized according to other parameters. To sum up, I suggest that WFR's should be divided into morphologically- phonologically- and syntactically- conditioned rules.

The WFR's that include regular derivation and compounding in Scalise's model will belong to what I am now calling [-phon] derivation and the evaluative rules will belong to the [+ phon] derivation in (18).

(18)



(Scalise's model revised)

Spanish morphology has other cases of [+ phon.] derivation besides the diminutive, augmentative and pejorative affixes. There are: hypocoristics, shortened forms, superlatives (*-ísimo, -érrimo* . . .), augmentative reduplication (*re-, requete-* . . .) and, perhaps, the controversial interfixes or "antihiatus" consonants (Malkiel 1958). They are all morphophonological processes of infixation, reduplication and copying, shortening and haplology, of a type, always considered problematic in a concatenative morphology.

We can predict, then, that if a morphological process is of the non-concatenative type, it will be phonologically conditioned.

What seem to be the most relevant features of the phonological word-formation rules? I put forward the following:

1) While typical derivational rules are limited to an individual level or stratum, with their affixes attached to any morphological layer, so long as its subcategorization frame is fulfilled, phonological derivation rules are not assigned to a precise morphological level, but are allowed to range over several strata. If an affix is attached, its frame will be defined in terms of a constituent determined phonologically, but not morphologically.

They are "terminal" derivative affixes, attached when the word, as a phonological unit, is already built up.

2) [- phon] rules may, or may not convey a morphophonological operation, as for example in the case of zero suffixation.

[+ phon] rules must necessarily induce a morphophonological change, although not necessarily through affixation, as for example in the case of shortened forms.

3) [- phon] rules are not reduplicated; the well-known examples with prefixes *re-* or *anti-* are cases of recursivity.

[+ phon] rules admit optional reduplication, resulting in intensification which does not affect the denotative meaning of the word.

4) [- phon] rules do not depend on the syllabic structure of the word. The position of the accent in the word base is not relevant, either.

[+ phon] rules are particularly sensitive to syllabic structure and the position of the accent.

5) [- phon] rules of derivation are disjunctively ordered in relation to each other. However, with respect to other [+ phon] derivation rules,

[+ phon] rules are not.

6) [- phon] rules are highly idiosyncratic and limited as far as productivity is concerned. There is no constant, predictable semantic result.

[+ phon] rules are much more productive and uniform; their semantic result is fixed and predictable, usually affecting the connotative meaning.

7) [- phon] rules apply only to bases that exhibit a concrete category label. Thus, they conform to the Unitary Base Hypothesis (UBH) (Aronoff 1976).

[+ phon] derivation rules, on the contrary, are indifferent to the syntactic category of the base.

8) [- phon] rules are governed by morphological well-formation conditions such as the "adjacency condition" or the "atom condition", and the more general condition of the UBH or the "Bracketing Erasure Convention".

[+ phon] derivation rules are not subject to well-formation conditions of a morphological character.

The contradictions I have pointed out in relation to Spanish diminutivization can, in my opinion, be explained if evaluative affixation is understood to be a word-formation process of a phonological nature.

Like the rules of the phonological component, phonological derivation rules have access only to information from the immediately adjacent cycle. This explains why the *-it-* infix, inserted inside the second constituent of the word, cannot change the gender marker (see examples in (14)): it does not see the next cycle, corresponding to the compound noun.

The fact that the evaluative affixation is not dependent on any type of morphological level-ordering, nor subject to morphological well-formation conditions like the rest of the derivation rules, explains the apparent contradiction of the examples in (13): the *-it-* infix is inserted into the word when it is already built up and, although contiguous to the second constituent, refers to the entire compound.

Finally, the fact that the evaluative rules, like all [+ phon] derivation rules, have access only to the internal phonological structure of the word, not to its morphological structure, would explain the apparent anomaly in (16).

References

Anderson, S.: Where's Morphology? In: Linguistic Inquiry 13 (1982), 571–612.
Aronoff, M.: Word Formation in Generative Grammar. The MIT Press, Cambridge, Ma. 1976.
Cabré, M. and Rigau, G.: Lexicologia i Semantica. Enciclopèdia Catalana, Barcelona 1986.
Contreras, H.: Spanish Exocentric Compounds. In: Nuessel, F. (ed.): Current Issues in Hispanic Phonology and Morphology. IULC, Bloomington 1985.
Hammond, M.: Level Ordering, Inflection and the Right-hand Head Rule. MIT Working Papers in Linguistics 7 (1984), 33–52.
Jaeggli, O.: Spanish Diminutives. In: Nuessel, F. (ed.): Contemporary Studies in Romance Languages. IULC, Bloomington 1980, 142–158.
Kiparsky, P.: Lexical Morphology and Phonology. In: Linguistics in the Morning Calm. Hanshin Publishing Co., Seoul 1982, 3–91.
Lloyd, P.: Verb-Complement Compounds in Spanish. Niemeyer, Tübingen 1968.
Malkiel, Y.: Los Interfijos Hispánicos. Problemas de Lingüística Histórica y Estructural. In: Miscelánea Homenaje a A. Martinet (II). 1958, 107–199.
Scalise, S.: Morfologia Lessicale. Clesp, Padova 1983.
Scalise, S.: Generative Morphology. Foris, Dordrecht 1984.
Siegel, D.: The adjacency condition and the theory of morphology. In: NELS 8 (1977), 189–197.
Williams, E.: On the Notions "Lexically Related" and "Head of a Word". In: Linguistic Inquiry 2 (1981), 245–274.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 245–260 (1986)

PHONOLOGICAL EVIDENCE FOR LEVEL ORDERING IN ITALIAN WORD FORMATION*

IRENE VOGEL

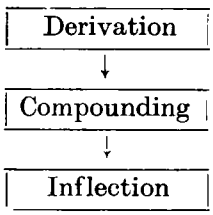
(University of Delaware, USA)

1. Introduction

It has often been noticed that morphemes do not combine freely and in any order to yield complex words. Instead, there appear to be strong constraints, language specific and perhaps also universal, on the way in which morphemes may be combined. One attempt to formalize such restrictions is the Level Ordering Hypothesis proposed by Siegel (1974). According to this proposal, the morphological component of the grammar consists of a number of ordered blocks of rules that carry out morphological operations; the order of occurrence of different morphemes is determined (at least in part) by the ordering of the blocks of rules and which types of rules are contained in the different blocks. Thus, the observation that in English the so-called Class I affixes appear before the Class II affixes is accounted for by having Class I affixation carried out by the first block of rules and Class II affixation by the second block of rules. In subsequent proposals, the original Level Ordering Hypothesis is expanded, as the Extended Level Ordering Hypothesis, to cover cases of compounding (cf. Allen 1975) and inflection (cf. Scalise 1984). Support for level ordering in morphology is adduced not only on the grounds that it accounts for a number of generalizations about the distribution of different types of morphemes, but also on the grounds that it accounts for certain phonological and semantic properties of affixes (cf. Pesetsky 1979), an idea that has been further developed within the framework of lexical phonology (cf. among others, Kiparsky 1983; Mohanan 1982; Halle and Mohanan 1985). When compounding and inflection are considered along with derivation, the basic ordering of the levels is generally taken to be that in (1), with language specific options such as more than one level of derivation as proposed for English, and more than one level of compounding as proposed for Malayalam (cf. Mohanan 1982).

* I would like to thank Marina Nespor and Sergio Scalise for their helpful comments on an earlier version of this paper.

(1)



This paper addresses the issue of level ordering in Italian. First, the question of the order of occurrence of different types of morphemes is examined. It will be shown that there are, in fact, generalizations about the structure of complex words that can be captured by a level ordering model of the lexicon of Italian, though the order is not that given above in (1). The rest of the paper is devoted to an examination of a number of phonological rules of Italian. Specifically, the domains in which these rules apply will be investigated with the goal of determining whether or not there is independent phonological evidence for the level ordering proposed on the basis of structural morphological considerations. It will be demonstrated that there is, in fact phonological support for lexical levels, though in one area morphological and phonological considerations lead to different conclusions about the composition of the levels.

2. Morphological ordering of levels

The ordering of derivation before both compounding and inflection seen in (1) has been proposed for Italian (cf. Scalise 1984) (cf. among others Kiparsky 1982; Mohanan 1982). This accounts for the fact that we find derived words within compounds in Italian, but we do not find compounds that have been further derived, as illustrated in (2) and (3), respectively.[1]

(2) (a) *formazione base* 'basic preparation' (from: *formare* 'to form')
(b) *città dormitorio* 'bedroom community' (from: *dormire* 'to sleep')
(cf. Scalise, 1984, 121)

(3) (a) *aspira polverista (from: *aspira polvere* 'vacuum cleaner')
(b) *banconotaio (from: *banconota* 'bank note')

It should be noted that the suffixes in (3), *-ista* and *-aio*, normally attach to nouns to form derived words with the general meaning of 'someone who has something to do with X (i.e. the noun)'. The difference between the gramma-

[1] There are several apparent exceptions to the generalization expressed here. For example, we do find words such as *croce rossina* 'Red Cross nurse' derived from *Croce Rossa* 'Red Cross'. As Scalise (1984) points out, however, the relevant cases are all instances of lexicalized compounds which are derived as any other, (noncompound) lexical item would be.

ticality of *macchinista* 'machinist' (from: *macchina* 'machine') and *libraio* 'book dealer' (from: *libro* 'book') on the one hand, and the ungrammaticality of the examples in (3) on the other hand, is the fact that the former are derived from a simple base while the latter are derived from a compound base.

That derivation applies before inflection, too, is seen by the fact that inflectional affixes always follow derivational affixes, no matter how many derivational affixes there are, as illustrated in (4), where the inflectional affixes are the final vowels.[2]

(4) (a)	*forma*	'form (sg)'
	forme	'form (pl)'
(b)	*formale*	'formal (sg)'
	formali	'formal (pl)'
(c)	*formalismo*	'formalism (sg)'
	formalismi	'formalism (pl)'

According to the ordering in (1) we would expect, furthermore, that compounding precedes inflection, that is, that inflectional affixes appear external, rather than internal, to compounds. This is, in fact, the ordering that is accepted for Italian by Scalise (1984). Consider, however, compounds such as those in (5) which have the structure Verb + Noun, where the noun is the object of the verb.

(5) (a)	*lava piatti*	'dish washer'	(*piatti* 'dishes')
(b)	*apri bottiglie*	'bottle opener'	(*bottiglie* 'bottles')
(c)	*gira dischi*	'record player'	(*dischi* 'records')

Notice that in the compounds in (5), which are representative of a very productive compound formation rule of Italian, the noun is plural, despite the fact that the entire compound is singular.[3] That is, despite the plural morpheme

[2] The fact that the adverb forming suffix *-mente* attaches to a form with the final vowel *a* and not *o* (e.g. *certamente* 'surely' vs. **certomente*) is sometimes taken as an argument that inflection may occur internal to derivation. While the forms in question come from Latin forms which were, in fact, feminine, there is no motivation for considering them feminine in modern Italian, although *a* is the vowel that usually indicates feminine gender. Instead, these forms can better be handled by an allomorphy rule that changes *o* to *a* before *-mente*, that is, by the same type of mechanism that is used to handle another phonological rule that applies in relation to *-mente*, the deletion of the final *e* of a base if it is preceded by a sequence of V + Liquid, as in *generale* + *mente* → *generalmente* 'generally'.

[3] While Verb + Noun compounds most commonly contain a plural noun, there are some that seem to allow either a singular or a plural noun (e.g. *segna libro/segna libri* 'book mark', where *libro* = book, *libri* = books), and others that typically take a singular noun. Some of these latter compounds seem to have a singular for rather idiosyncratic reasons (a point we will return to below), while others take a singular for very specific, predictable, reasons. For example, we typically find a singular in compounds in which the

-i in *lava piatti*, the compound exhibits singular agreement with all the other words in a sentence:

(6) *È una lava piatti costosa.* '(It) is an expensive dish washer.'

If inflection applies after compounding as suggested in (1), compounds such as those in (5) would be constructed as in (7).

(7)

[piatto]	
[[lava] [piatto]]	compounding
[[lava piatto] i]$_{pl}$	inflection
[lava piatti]	'dish washers'

The problem with (7) is that the plural marker *-i* is attached to the entire compound and we are thus unable to get the desired meaning of 'only one dishwasher'. If the order of the operations is reversed, however, and we allow inflection to apply before compounding, the construction of the compounds in question would proceed as in (8).

(8)

[piatto]	
[[piatto] i]$_{pl}$	inflection
[[lava] [piatti]]	compounding
[lava piatti]	'dish washer'

The advantage of (8) over (7) is that (8) captures the fact that the plural suffix is actually internal to the compound, despite its position as the rightmost morpheme.

It should be noted that Italian is not the only language that allows inflection within compounds. An analogous situation arises, for example, in Spanish which also has a productive rule of compound formation which creates words of the structure Verb + Noun, in which the noun object of the verb is usually plural (e.g. *toca discos* 'record player', where *discos* = 'records') (cf. Varela 1986). In English, too, we find inflection within compounds as in *systems analyst, parks commissioner*. In all these cases, if inflection only applies after

noun is a mass noun and thus cannot normally be pluralized (e.g. *aspira polvere* 'vacuum cleaner' (lit. 'suck dust'), where the noun *polvere* 'dust' is singular). In addition, the presence of a singular form is often predictable on the basis of pragmatic considerations, as in cases in which the noun in question refers to a unique entity in the world, as in the hypothetical example suggested by W. U. Dressler of a machine for watching the sun, a *guarda sole*, where *sole* 'sun' is singular since it is assumed (at least in nonscientific circles), that there is only one sun. The same is true also for cases in which the noun must, for some other reason, refer to a single object, as in the case of *ficcanaso* 'busybody' (lit. 'shove nose'), where *naso* 'nose' is singular for obvious anatomical reasons. The plural morpheme, *-i*, in *ficcanasi* 'busybodies' refers to the entire compound, that is, giving the meaning of more than one busybody, not the meaning of a person who puts more than one nose where it does not belong.

compounding, we have no way of getting the plural meaning of one of the component words of the compound but not the plural of the whole compound. Of course, is it often necessary to make the entire compound plural as well. If inflection has already taken place before compounding, however, the problem arises as to how to get, for example, *lava piatti* with the meaning of 'dish washers', or *systems analysts* with two plural markers. It should be noted that the surface form *lava piatti* also represents the plural of the entire compound, as illustrated by the plural agreement in (9), in contrast with the singular agreement seen above in (6).

(9) *Sono delle lava piatti costose.* '(They) are some expensive dish washers.'

One possibility would be to allow a loop of the type proposed by Mohanan (1982) to take an item back to inflection after compounding for a second application of inflection. Loops, however, greatly weaken the model since they often allow for many more possibilities than are otherwise required by a given language what I propose instead to treat the compound internal application of inflection separately from the inflection of the entire compound. That is, I propose that the former be treated lexically by word formation rules and the latter syntactically.[4] The issue of whether or not one of the components of a compound is plural is not, after all, what is relevant for syntax. In the case of agreement, for example what matters is only whether the entire compound is plural, in keeping with the idea that syntax is not sensitive to the internal structure of words. Of course, if we maintain that there is a distinction between morphological and syntactic types of inflection, we predict that there ought to be some substantive difference between the two. For example, we might expect that the former, but not the latter, exhibits somewhat idiosyncratic patterns. This is, in fact, what we find. Specifically, despite Italian speakers' explanation that *lava piatti* contains the plural form *piatti* because it would not make sense for there to be a machine that washes only one dish (*piatto*), the use of the plural often seems somewhat arbitrary. The compound *apri bottiglie* 'bottle opener' also contains a plural even though it can open only one bottle at a time. If the crucial fact is that the object in question can be used for a multitude of bottles on separate occasions but only one at a given time, we cannot explain why another compound *porta badiera* 'flag carrier' contains a singular noun *bandiera* 'flag'. At any one time, it is true, a flag carrier carriers one flag, but like the bottle opener, on different occasions it might carry different flags. Furthermore, if *lava piatti* contains a plural noun because the machine is made to wash many dishes at once, why does the corresponding compound in English, *dish washer*, contain a singular noun? What these examples show is that the choice of

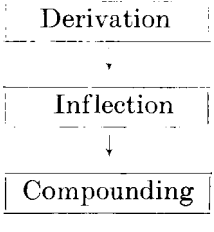
[4] See Hammond (1984) for a similar proposal for English.

whether the compounds in question take a plural or singular noun is not totally predictable. While there appears to be a language specific parameter involved which accounts for the fact that certain compounds normally take a plural noun in Italian where the corresponding compounds take a singular in English, there nevertheless also appears to be some idiosyncracy involved, as in cases such as *apri bottiglie* vs. *porta bandiera*. In contrast with the somewhat arbitrary nature of this compound internal type of inflection, the inflection that is syntactically determined does not exhibit such idiosyncacy. As was seen above in (6) and (9), repsectively, if *lava piatti* is singular, then the verb, article and adjective in the sentence must also be singular, and if *lava piatti* is plural, then the other words must also be plural.

It should be noted, however, that the proposal to treat word internal inflection separately from the type of inflection that is more syntactic in nature is not without its own difficulties. The major problem is that, despite certain differences between the two cases of inflection, the morphemes involved are identical, and treating the two cases separately makes their similarity seem fortuitous. Nevertheless, as we have seen, the position of certain inflectional affixes internal to compounds, as well as the fact that these instances of inflection are irrelevant to syntax, seems to indicate that, in addition to the regular, syntactic, type of inflection a lexical operation of inflection is also needed, one which allows for idiosyncracy, as do other lexical rules. This type of inflection must thus be considered part of the word formation component of a grammar, along with derivation and compounding. Anderson's (1982) claim that inflectional structure is developed in a different way from derivational structure woulds thus distinguish between the "internal" type of inflection, which would be developed more like derivational structure, and the "external" type of inflection, which would be developed in a different way. The theory of morphology will ultimately have to account for the relation between the two types of inflectional phenomena. We will not, however, deal with this problem further here; instead, we will limit our attention to the lexical operation of inflection and its interaction with other lexical morphological phenomena.

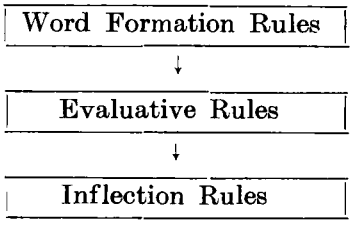
On the basis of the arguments considered above, let us now change the order of the lexical levels from that seen in (1) to that given in (10), at least for Italian, and perhaps more generally. This new order now gives the derivation seen above in (8).

(10)



There is still another problem in relation to the ordering of the lexical levels that must be addressed here. Scalise (1984) argues on the basis of a number of characteristics of the evaluative suffixes of Italian (i.e. diminutive augmentative, etc.) that these affixes cannot be considered either derivational or inflectional morphemes. Instead, they constitute a class of their own, and are associated with a distinct block of rules in the lexicon which is ordered between the blocks corresponding to the word formation rules and the inflectional rules, as seen below (cf. Scalise 1984, 133).

(11)



According to this schema it would seem that the evaluative rules should follow both compounding and derivation since both are considered to be types of word formation rules. While Scalise does not address the question of the relative ordering of compounding and the evaluative rules, this issue becomes crucial here since we are proposing to reorder inflection and compounding with respect to the order proposed by Scalise. If we move inflection up from the last level to the second level, do the evaluative rules remain where they are and thus become the last level, or do they move up along with inflection to the position before compounding? As it turns out, compounds in Italian do not normally undergo evaluative rules. That is, we do not typically find cases of a compound to which an evaluative suffix has been added. We cannot, for example, form the diminutive of *lava piatti* by adding the suffix *-ino* or the augmentative of *apri bottiglie* by adding the suffix *-one*, as illustrated by the ungrammaticality of the forms in (12).[5]

(12) (a) *lava piatti* + *ino* → **lava piattino*
(b) *apri bottiglie* + *one* → **apri bottiglione*

Thus, since compounding is, in fact, a word formation rule, it is questionable whether the evaluative rules should be ordered after the entire block of word formation rules, as in (11). If, on the other hand, the evaluative rules are moved before compounding, we predict that evaluative affixes may appear

[5] Spanish seems to differ from Italian with respect to this point since in Spanish it is possible to add evaluative affixes to analogous Verb + Noun compounds. For example, it is possible to make the diminutive of *toca discos* 'record player' by adding the infix *-it-* : *toca disquitos* 'small record player' (cf. Varela 1986).

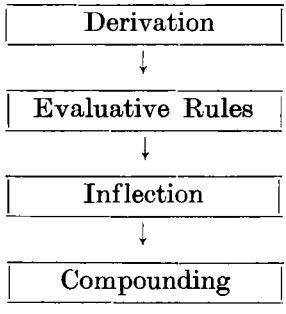
within compounds. As it turns out, however, we do not typically find evaluative affixes within compounds either, though the reason for this might be more pragmatic that structural. That is, it does not seem all that likely that there would be a machine, for example, that washes only small dishes, or an object that opens only large bottles, and hence the unacceptability of the forms in (13). Note that the nouns are given in their plural form in keeping with the typical pattern for such compounds.

(13) (a) *lava* + *piattini* → **lava piattini* (cf. *piattini* 'small dishes')
(b) *apri* + *bottiglioni* → **apri bottiglioni* (cf. *bottiglioni* 'large bottles')

Thus far, however, we have no compelling arguments either in favor of the ordering in (11) or against it.

Let us now consider the relationship between the evaluative rules and inflection since we are proposing to move inflection with respect to compounding. In fact, it turns out that evaluative affixation must take place before inflection since evaluative affixes always appear internal to inflectional affixes (e.g. *piattino/piattini* 'small dish (sg)/small dish (pl)'; *bottiglione/bottiglioni* 'large bottle (sg)/large bottle (pl)'). We must thus maintain the ordering of evaluative rules before inflection rules as proposed by Scalise (1984), which in the present proposal requires that the block of evaluative rules be moved before compounding along with the movement of inflection. The resulting organization of lexical levels in Italian must thus be the one in (14), not the one in (11). In the next section we will examine this proposal further to see if there is any independent motivation for such a model.

(14)



3. Phonological evidence for lexical levels

As we have just seen, the organization of four types of morphological rules as a sequence of blocks of rules accounts for a number of distributional patterns of morphemes within complex words in Italian. The ordering proposed here differs somewhat from the order proposed for several other languages,

and more importantly for our purposes, from the order previously proposed for Italian. Since one of the motivations for the original proposals of level ordering in English morphology was the fact that the proposed model also accounted for a number of phonological phenomena, we will now examine some rules of Italian phonology to see whether they provide any support for lexical levels, and if so, whether the organization proposed above is correct. The type of rules we will be concerned with here are phonological rules that apply in relation to morphological structure, but not in relation to specific morphemes.[6] The latter type of rules, allomorphy and truncation rules in Aronoff's (1976) terms, cannot provide crucial evidence in relation to lexical levels and their ordering since they only apply in relation to a single word formation rule (e.g. $e \rightarrow \emptyset / V \begin{bmatrix} +\text{cons} \\ +\text{son} \\ -\text{nas} \end{bmatrix} \_\_ + \text{mente}$:

liberale + *mente* → *liberalmente* 'liberally'; *regolare* + *mente* → *regolarmente* 'regularly'). That is, the very restricted contexts of such rules do not allow us to make more general statements about the domains of application of phonological rules and the interactions among them.

Among the phonological rules that apply in relation to morphological structure are two stress rules. The first one, the Word Stress Rule (WSR), assigns stress to one of the last three syllables of a word.[7] This rule, which was shown to apply cyclically by Vogel and Scalise (1982), applies in the same way to derived and inflected words, as well as to words with evaluative affixes, as seen in (15a)—(15c), respectively. Only primary stress, indicated with an acute accent (') will be considered here.

(15) (a)	*favorévole*	'favorable'	(from: *favore* + *evole*)
(b)	*elefánti*	'elephants'	(from: *elefante* + *i*)
(c)	*cappellíno*	'small hat'	(from: *cappello* + *ino*)

The second rule, the Compound Stress Rule (CSR), assigns stress to the syllable bearing the primary stress in the rightmost member of a compound, with the concomitant weakening of the primary stress of the other members of the compound, as seen in (16).

[6] This is not to say that the rules in question may not apply differentially to specific morphemes. In fact, one of the characteristics of lexical phonological rules is that they may have exceptions, as opposed to post-lexical rules which are said to apply across the board without exceptions. The point is that the phonological rules we will consider here do not apply uniquely in relation to specific morphemes.

[7] There is a small set of verbs which have stress on the fourth to last syllable of the third person plural forms of the present indicative and the present subjunctive (e.g. *teléfonano* '(they) call', *teléfonino* '(that they) call). Since these exceptions are so circumscribed, they do not constitute a serious problem for the generalization about the position of stress in Italian words.

(16) (a) *apri bottíglie* 'bottle opener' (from: *ápri* + *bottíglie*)
(b) *pelle róssa* 'red skin' (from: *pélle* + *róssa*)
(c) *tosta páne* '(bread) toaster' (from: *tósta* + *páne*)

Let us now consider four segmental rules that also apply in relation to morphological structure. The first rule is Vowel Deletion (VD), the rule that deletes an unstressed vowel when it is followed by another vowel in the following morpheme (cf. Scalise 1983), as formulated in (17). The symbol ']' is used in this and subsequent rules to indicate the limit of a morpheme without specifying anything about the nature of the morphemes involved since one of the advantages of a theory that incorporates distinct levels in the lexicon that can also account for the domains of application of phonological rules is that the specification of different types of boundaries is no longer necessary.[8]

(17) Vowel Deletion (VD)
$\breve{V} \rightarrow \emptyset / \_\_] V$

This rule applies to derived and inflected words, as well as to words with evaluative affixes, as illustrated in (18a)—(18c), respectively.

(18) (a) *paură — oso → pauroso* 'fearful'
(cf. *virtú — oso → virtuoso/*virtoso* 'virtuous')
(b) *paură — e → paure* 'fears'
(c) *piattŏ — ino → piattino* 'small dish'

Notice that this rule does not apply morpheme internally: the first unstressed *a* in (18a, b) is not deleted before the following vowel. Furthermore, VD does not apply in the same way to compounds, where it is often optional and subject to a number of additional constraints (e.g. *porta ombrelli/port ombrelli* 'umbrella rack') (cf. Vogel and Scalise 1982).[9]

Since VD makes crucial reference to stress, and since the assignment of stress is cyclic, an obvious question at this point is whether VD is also cyclic. Consider the words *algébrico* (from: *algebra — ico* 'algebraic' and *virtuóso* (from: *virtú — oso*) 'virtuous'. If VD applies before stress assignment both the final *a* of *algebra* and the final *u* of *virtú* will be deleted, giving the incorrect result in the latter case. If stress assignment applies first, the *a* preceding *-ico* will be stressed, thus exempting it from VD; the *o* of *oso* will also be stressed, leaving the preseding *u* available for the application of VD. The result is

[8] Stress is indicated here and below as a diacritic feature on the vowel: '$\breve{V}$' = unstressed, '$\acute{V}$' = stressed. This is done for reasons of graphic simplicity; it does not mean that stress should necessarily be viewed as a feature associated with individual vowels.

[9] As is often the case with lexical phonological rules, there are some exceptions to VD (e.g. *veritá + iero → veritiero* 'truthful').

incorrect in both cases. This problem can be resolved easily, however, by applying both the WSR and VD cyclically, as illustrated in (19).[10]

(19) (a)	Cycle 1	*algebra*	*virtu*
	WSR	*á*	*ú*
(b)	Cycle 2	*álgebra ico*	*virtú oso*
	VD	*ø*	—
	WSR	*é*	*ó*
		algébrico	virtuóso

The next two rules we will examine, Vowel Lengthening (VL) and Vowel Raising (VR), also depend crucially on stress. As the rule in (20) shows, VL lengthens a stressed vowel in a nonfinal open syllable. VR, as seen in (21), raises [ɛ] and [ɔ] to [e] and [o], respectively, when they are not stressed.

(20) Vowel Lengthening (VL)

$\acute{V} \rightarrow [+\text{long}] / \_\_ \$\ C_0V\ldots]$[11]

(21) Vowel Raising (VR)

$$\begin{bmatrix} +\text{ mid} \\ +\text{ low} \end{bmatrix} \rightarrow [\text{-low}] / \left[\overline{\text{-stress}}\right]$$

Both VL and VR apply in relation to derivation, inflection and evaluative affixation, as seen in (22) and (23).

(22) (a)	*lava — abile → lav[á:]bile*	'washable'
(b)	*lava — vamo → lavav[á:]mo*[12]	'(we) were washing'
(c)	*bravo — ino → brav[í:]no*	'good (dim)'
(23) (a)	*tosta — tore → t[o]statóra*	'roaster'
	(cf. *t[ɔ]sta* '(he) roasts')	
(b)	*tosta — vamo → t[o]stavámo*	'(we) were roasting'
(c)	*cosa — ina → c[o]sína*	'thing (dim)'
	(cf. *c[ɔ̄]sa* 'thing')	

It should be noted that VL and VR must apply separately to the two members of a compound, before the application of the CSR which places stress on the final member of the compound, automatically eliminating the primary stress

[10] Note that along with each application of the WSR the primary stress assigned on the previous cycle is automatically weakened (cf. Vogel and Scalise 1982).

[11] The symbol '$' is used as a type of shorthand notation here; it is not meant to imply that syllable structure should be seen in terms of boundaries inserted between segments.

[12] Alternatively, the inflectional suffix may be taken to be *-avamo*, in which case VD will apply to delete the stem *a*. As far as the present point is concerned, however, both analyses yield the same results.

of any other members. That is, we find a lengthened vowel in both members of the compound in (24a); it is not just the last stressed vowel of the entire compound that is lengthened as we would expect if VL applied after the CSR had eliminated all but the final primary stress. Similarly, if the CSR had caused the stress on *reggi* in (24b) to be eliminated before VR applied, we would expect to find a raised vowel, [e] instead of [ɛ] in this word.

(24) (a) *capo — stazione* → *c*[á:]*po stazi*[ó:]*ne* 'station master'
(b) *reggi — petto* → *r*[é]*ggi p*[é]*tto* 'brasière'

It was seen above that VD, which crucially makes reference to stress, must apply cyclically. Since both VL and VR also crucially make reference to stress, the question arises as to whether they, too, must apply cyclically. Let us consider VL first. If this rule applies cyclically it will lengthen a vowel in the proper segmental context every time the WSR applies. In a complex word in which a given number of affixes have been added, we would thus expect there to be potentially the same number of lengthened vowels, as illustrated in (25). This is not correct, however; there is at most only one long vowel per word.

(25)			
	Cycle 1	*noia*	'boredom'
	WSR	*ó*	
	VL	*ó:*	
	Cycle 2	*nó: ia oso*	'boring'
	VD	Ø	
	WSR	*ó*	
	VL	*ó:*	
	Cycle 3	*no: io:so ino*	'boring (dim)'
	VD	Ø	
	WSR	*i*	
	VL	*i:*	
	*[*no:io:sí:no*]	(cf. [noiosí:no])	

One solution would be to introduce another rule, one which deletes all the long vowels except the final one, though at this point we might question the desirability of lengthening all the vowels in the first place. In fact, a simpler and more insightful solution is to allow VL to apply only once, rather than cyclically. In this way, VL applies only to the final output of the various morphological operations (except compounding).

Let us now consider what happens if VR applies cyclically after stress is assigned:

(26)	Cycle 1	*alg*[ɛ]*bra*	'algebra'
	WSR	*á*	
	VR	*e*	
	Cycle 2	*álg*[e]*bra ico*	'algebraic'
	VD	Ø	
	WSR	*é*	
	VR	-------	
		**alg*[é]*brico*	(cf. *alg*[*é*]*brico*)

The problem here is that if VR raises [ɛ] to [e] after stress has been assigned on the first cycle there is no way to get the [ɛ̆] back in the derived form in which the vowel is stressed, since not all instances of stressed [e] become [ɛ]. The solution to this problem too is to allow VR to apply only once, after all the morphological operations (except compounding) have applied. Since we want VL and VR to apply as late as possible, but before compounding, we can state the domain of their application in terms of lexical levels as being the level just before compounding, that is, inflection, according to the order proposed above.

Finally, let us consider a rule we will call Consonant Lengthening (CL). This rule lengthens the initial consonant of the second member of a compound when the first member ends in a stressed vowel,[13] as stated in (27) and illustrated in (28).

(27) Consonant Lengthening (CL)

$$C \rightarrow [+\text{long}] / \begin{bmatrix} V \\ +\text{stress} \end{bmatrix}]\_\_ [+\text{son}]$$

(28) (a) *caffé — latte → caffé* [l:]*atte* 'coffee with milk'

(b) *cittá — dormitorio → cittá* [d:]*ormitorio* 'bedroom community'

Table 1

Domains of application of phonological rules

	Derivation	Evaluative Affixation	Inflection	Compounding
Word Stress Rule (WSR)	+	+	+	−
Vowel Deletion (VD)	+	+	+	−
Vowel Lengthening (VL)	−	−	+	−
Vowel Raising (VR)	−	−	+	−
Compound Stress Rule (CSR)	−	−	−	+
Consonant Lengthening (CL)	−	−	−	+

[13] Consonant Lengthening in compounds actually falls under a more general rule of consonant lengthening, Raddoppiamento Sintattico. The general form of the rule applies to a consonant at the beginning of a (phonological) word when it is preceded by a word ending in a stressed vowel whenever this sequence arises within the domain of the phonological phrase. It thus applies to the members of a compound as it does to words in a phrase (cf. Nespor 1984; Nespor and Vogel 1986).

What our examination of the two stress rules and the four segmental rules of Italian considered in this section shows is that not all of the rules apply in relation to all morphological operations. The following table presents an overview of the rules and their domains of application.[14] The symbols '+' and '—' mean that the rule applies and that it does not apply, respectively.

Let us now consider what the distribution of the six phonological rules considered here says about the organization of the Italian lexicon. First of all is there phonological evidence for the levels proposed on the basis of the morphological criteria discussed above in Section 2? In general, the answer to this question is affirmative. That is, compounding is distinguished phonologically from all the other morphological operations by the fact that it is the only domain for CSR and CL, while the rules that apply in other domains do not apply in relation to compounding. Inflection, on the one hand, and derivation and evaluative affixation, on the other hand, are also phonologically distinct from each other in that the former is the only domain for VL and VR. In other words, there are specific phonological rules that apply in relation to specific types of morphological operations. We thus have support for individual lexical levels corresponding to a) derivation (and evaluative affixation), b) inflection and c) compounding. Although Scalise (1984) provides morphological arguments for establishing a separate level for evaluative affixation, as mentioned above, no phonological evidence for such a level has been found. On phonological grounds, evaluative affixation must be considered part of the level of derivation.

Secondly, what do the phonological rules say about the relative order of the three lexical levels for which there is evidence? The crucial rules are WSR and VD. Compare the domains of application of these rules as ordered in (29a) and (29b).

(29) (a)

	Derivation Evaluative Affixation	Compounding	Inflection
WSR	+	—	+
VD	+	—	+

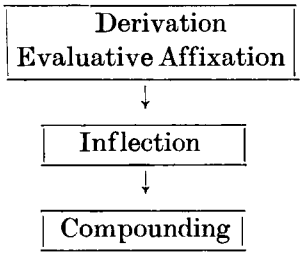
[14] Several of the rules analyzed here have also been analyzed within the framework of prosodic phonology (cf. Nespor 1984; Nespor and Vogel 1986). In particular, it is shown that the domain of the rules in question is the prosodic constituent referred to as the phonological word. The prosodic approach is not in conflict with the one presented here, and in fact, the two approaches complement each other in that the prosodic analysis specifies the phonological domain within which the rules apply and the lexical analysis accounts for which types of morphological structures are subject to particular phonological rules. See also Booij and Rubach (forthcoming) where it is argued that that in Polish, too, the phonological word plays a crucial role in the rules of lexical phonology.

(b)

	Derivation Evaluative Affixation	Inflection	Compounding
WSR	+	+	—
VD	+	+	—

In (29a), the ordering proposed by Scalise (1984) for Italian, and by other linguists for other languages, leads to a violation of the Continuous Stratum Hypothesis proposed by Mohanan (1982), according to which, if a given phonological rule applies to more than one lexical level (stratum in the terms of lexical phonology), the levels in question must be adjacent. That is, since WSR and VD do not apply to compounding, if compounding is ordered between the other two levels to which the rules do apply, the domains of application of the rules are not adjacent. The alternative ordering proposed in this paper does not lead to a violation of the Continuous Stratum Hypothesis, as (29b) shows. We thus find phonological support not only for three lexical levels, but also for their reordering such that compounding, not inflection, is the last stratum, as respresented in (30).

(30)



4. Conclusions

In this paper, we have examined the question of the existence and organization of lexical levels in Italian on the basis of both morphological and phonological phenomena. Specifically, it was shown that there seems to be morphological evidence for four levels, corresponding to the operations of derivation, evaluative affixation, inflection and compounding. Morphological considerations having to do with the distribution of different types of morphemes, it was seen further, led us to propose an ordering that is different from that previously proposed for Italian. Examination of a set of phonological rules, however, provided support for only three lexical levels; derivation and evaluative affixation could not be distinguished phonologically. This result points up a problem related to lexical levels in general: what are the criteria

on which we establish lexical levels? Presumably when we have both morphological and phonological evidence for a level the case is the strongest. But what do we do if there is a conflict between the number of levels that can be justified phonologically and the number of levels that can be justified morphologically, as was seen here in relation to Italian? In general, what seems to happen is that morphologists tend to accept the results provided by the morphological analysis, as illustrated by Scalise's (1984) proposal to establish the additional level of evaluative affixation, while phonologists seem to favor the results provided by the phonological analysis, as illustrated by some recent proposals within the framework of lexical phonology to have only two lexical strata on the basis of a purely phonological criterion, cyclic vs. noncyclic rule application (cf. Kiparsky 1983; Booij and Rubach, to appear). I have no solution to this problem now but certainly in order to make progress in our understanding of the organization of the lexicon, the relationship between morphological and phonological criteria and their respective roles in determining lexical structure will have to be reconciled. Finally, leaving aside the problem of the evaluative suffixes, it was shown that there is evidence from both morphology and phonology for reordering inflection and compounding so that the level associated with inflection precedes the level associated with compounding.

References

Allen, M. R.: Morphological Investigations. Ph. D. Diss., University of Connecticut 1975.
Anderson, S. R.: Where's morphology? In: Linguistic Inquiry 13 (1982), 571–612.
Aoronoff, M.: Word Formation in Generative Grammar. MIT Press, Cambridge, Mass. 1976.
Booij, G. and Rubach, J.: Postcyclic versus postlexical rules in lexical phonology. In: Linguistic Inquiry 18 (1987), 1–44..
Halle, M. and Mohanan, K. P.: Segmental phonology of modern English. In: Linguistic Inquiry 16 (1985).
Hammond, M.: Level ordering, inflection and the righthand head rule. In: MIT Working Papers in Linguistics 7 (1984), 33–52.
Kipasrky, P.: From cyclic to lexical phonology. In: Hulst, H.v.d. and Smith, N. (eds): The Structure of Phonological Representations: Part I. Foris, Dordrecht 1982.
Kiparsky, P.: Some Consequences of Lexical Phonology. Manuscript, MIT 1983.
Mohanan, K. P.: Lexical Phonology. Ph. D. Diss., MIT (also IULC) 1982.
Nespor, M.: The phonological word in Italian. In: Hulst, H.v.d. and Smith, N. (eds): Advances in Non-Linear Phonology. Foris, Dordrecht 1984, 193–204.
Nespor, M. and Vogel, I.: Prosodic Phonology. Foris, Dordrecht 1986.
Pesetsky, D.: Russian Morphology and Lexical Theory. Manuscript, MIT 1979.
Scalise, S.: Morfologia Lessicale. CLESP, Padova 1983.
Scalise, S.: Generative Morphology. Foris, Dordrecht 1984.
Siegel, D.: Topics in English Morphology. Ph. D. Diss., MIT 1974.
Varela, S.: Inflectional and diminutive affixation inside Spanish noun compounds. Paper read at the Conference on Theoretical Approaches to Morphology, Veszprém, Hungary, May 1–4, 1986. Cf. this volume p. 235–244.
Vogel, I. and Scalise, S.: Secondary stress in Italian. In: Lingua 58 (1982), 213–242.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 261–271 (1986)

DERIVATION CONSTRUCTIONS AND ARGUMENT STRUCTURE

W. ZWANENBURG
(University of Utrecht, The Netherlands)

0. The central idea of this article is that derivation constructions of languages come in limited numbers and are to be defined, independently of individual affixes, in terms of X-bar structure and argument structure. In terms of X-bar structure we can define the categories and the order of the head and the complement of each — supposedly binary — derivation construction. In terms of argument structure we can define how the argument structure of the construction relates to that of the lexical morphemes which may serve as its bases.

To make this general idea a little more concrete, let us consider an example. One of the best known derivation constructions is that of action nouns like those in (1):

(1) *slippage, arrival, resistance, conversation, opening, confusion, amusement*

This construction is independent of individual affixes in that action nouns may contain any of a long series of suffixes, among which those given in (1). In terms of X-bar structure they contain a verbal base as complement and a nominal suffix as head. As to their argument structure, according to Williams (1981) this can be derived from that of the verbal base by means of internalization of the external argument: the subject of the verb as the external argument becomes an internal argument of the action noun, as in (2):

(2) *John opens the exhibition → the opening of the exhibition by John*

I want to show here on the example of English suffixation that the limited number of possible constructions is the result of the restrictions on X-bar structure and argument structure. X-bar structure allows only — supposedly binary — combinations of the categories N, A and V, the second constituent being an affix in the case of suffixation as opposed to prefixation, and either the first or the second constituent being the head. Argument

structure allows a limited number of arguments, a limited number of manipulations of argument structure in derivation, and, as developed in Chierchia (1985),a limited number of relevant roles for nouns, adjectives and verbs at the level of predication.

A by-product of my argumentation will be that it is necessary to modify Williams' notion of external argument as far as nouns are concerned. It is also necessary to add as possible manipulations of argument structure the creation of a 'possessive' argument as well morphological incorporation. Cf. for other proposals concerning increasing the number of possible manipulations Carrier-Duncan (1985) and Zubizarreta (1985).

I represent in (3)(a)-(c) what I think are globally the suffication constructions of English. With very little changes the same table could be used for languages like Dutch and German, or like French, Italian and Spanish. In sections **1** through **6**, I will discuss successively the constructions contained in (3)(a)-(c). The concluding section **8** will be preceded by a short section concerning the probable degree of generality of the regularities discussed in this paper.

(3) (a)

	N	A	V
N	*drop-let*	—	—
A	—	*green-ish*	—
V	—	—	*spark-le*

(b)

	N
N	*queen-ship*
A	*kind-ness*
V	*open-ing*

(c)

	N	A	V	V internalization
N 'possessive'	*villag-er*	*gloom-y*	*agon-ize*	*hammer-Ø*
N	—	*child-ish*	*head-Ø*	*mumm-ify*
A	*drunk-ard*	—	*idle-Ø*	*hard-en*
V	*sing-er*	*fidget-y*	—	*run-Ø*
$V_{externalization}$	*read-er* *slipp-er*	*handle-able* *reli-able*	*melt-Ø*	—

1. In (3)(a) one finds the non category changing suffix constructions: denominal nouns, de-adjectival adjectives and deverbal verbs. In English as well as in many other languages suffixes tend to be category changing and prefixes non category changing. So the constructions of (3)(a) are marked in this respect as compared to those of (b) and (c), which are category changing. I assume that their being non category changing is the result of their lexical base being, in terms of X-bar structure, the head of the construction, as opposed to the constructions in (3)(b)-(c), which have suffixes as heads.

It seems to me that the constructions of (3)(a) do not allow changes in argument structure of the kind proposed by Williams 1981, such as externalization of an internal argument or internalization of an external argument. At most they will present deletion of one or more argument positions. In which conditions such deletions exactly take place is a question which cannot be answered within the limits of this paper.

Let me just try to shortly indicate the difference between such deletions and the more drastic changes discussed by Williams. The agent noun construction of *flyer*, illustrated in (4)(a), presents such a drastic change in the form of morphological incorporation of the external argument of the verb (see section **6** below). This is an essential characteristic of the construction. But in this particular case the change is accompanied by the obligatory deletion of an internal argument, as (4)(b) shows.

(4) (a) *John flies* → *John is a flyer*
(b) *John flies in circles* —) **John is a flyer in circles*

This deletion is not characteristic of the agent noun construction as such, but can occur in particular manifestations of it as well as of other constructions.

What is important to notice is the fact that the non category changing constructions of (3)(a) extend systematically over all three possible categories: N, A and V. The analysis given here allows no other possibilities for non category changing suffix constructions.

2. In (3)(b) we find abstract nouns, again derived systematically on the basis of the three lexical categories N, A and V. In terms of X-bar structure they contain a nominal, adjectival or verbal base as complement and an (abstract) nominal affix as head. Williams (1981) also treats these constructions together as nominalizations of nouns, adjectives and verbs.

Williams characterizes the argument structure of these constructions as resulting from the argument structure of the base by internalization of the external argument:

(5) (a) *Mary is a queen* → *the queenship of Mary*
(b) *Mary is kind* → *the kindness of Mary*
(c) *Mary opens the book* → *the opening of the book by Mary*

Derived nouns of the class of *opening* are traditionally indicated by the term action noun, the other ones are sometimes called abstract nouns.

What the examples of (5) show is that the peculiarity of these nouns consists in the fact that they represent a way of indicating nominally a nominal,

adjectival or verbal predicate. To see what exactly is going on, it is useful to consider Chierchia (1985)'s formal semantic theory of properties (or propositional functions) and predication, and the role it attributes to nominal, adjectival and verbal expressions.

Chierchia formulates a Three Layers Hypothesis (429). This says that there are three layers of types (or formal semantic categories) in predication: individuals or arguments, propositional functions or predicates, and third-order functors (424). The role of individuals or arguments is normally performed by nominal expressions, that of propositional functions of predicates by adjectival and verbal expressions, and that of third-order functions by modifiers of different sorts such as adverbs. Two particularities of Chierchia's reasoning must be added which are essential for my purpose. First, nominal expressions, which normally perform the role of individuals (*the president*), can under certain conditions play that of propositional functions (*I want Fritz p r e s i d e n t*), i.e. nominal expressions can be used as predicates (433—434). And second, "properties are projected in two distinct logical roles: as predicates and as [individuals]", i.e. predicates can be nominalized (422). Chierchia's Three Layers Hypothesis says then about the formal semantic roles of the categories used in morphology: (i) Nominal expressions are normally arguments, but they can be used as predicates (*I. want Fritz p r e s i d e n t*). As arguments, they can be the result of the nominalization of a predicate. (ii) Adjectival and verbal expressions are used as predicates.

This theory gives us a satisfactory means to interpret the nouns of (3)(b) and even to predict a considerable part of their properties. First of all, it predicts that predicates must be nominalizable, as we find in these nouns. Second, it predicts that there are three categories of predicates, to wit adjectives and verbs in their normal role and nouns in their secondary role. This means that the abstract nouns of (3)(b) realize morphologically what Chierchia's Three Layers Hypothesis predicts in this respect, nothing less and nothing more.

In relation to the constructions discussed in this section, we can observe then once more that they extend systematically over the three possible categories N, A and V.

3. Let us consider next the constructions illustrated within the frame in (3) (c). In terms of X-bar structure we have here a nominal, adjectival or verbal base as complement and a nominal, adjectival or verbal suffix as head of the construction.

As for predication, these constructions can be described as the transposition of nominal, adjectival and verbal predicates into words of one of the other two categories:

(6) (a) *John is a child* → *John is childish*
(b) *John is the head of the troop* → *John heads the troop*
(c) *John is drunk* → (*John is a*) *drunkard*
(d) *John is idle* → *John idles*
(e) *John sings* → (*John is a*) *singer*
(f) *John fidgets* → *John is fidgety*

The presence of these six constructions and the absence of the other three theoretically possible ones within the frame of (3)(c) are predicted by Chierchia's Three Layers Hypothesis.

To see this, consider again what this theory predicts for nouns, adjectives and verbs: (i) Nouns are normally arguments, they can be used as predicates, and as arguments they can be the result of nominalization of predicates. (ii) Adjectives and verbs are predicates.

Nominalization of nominal, adjectival and verbal predicates has been accounted for in (3)(b), so we do not have to discuss nominalization any more. Apart from that, for the simple and basic exchanges between categories, represented within the frame in (3)(c), this is what remains and what turns out to be realized according to the examples of (6): the transposition of nominal, adjectival and verbal predicates into adjectives and verbs with their necessarily predicative role, and furthermore the transposition of adjectivel and verbal predicates into nouns with their normal role of arguments (*drunkard* and *singer*). Predicative use of de-adjectival and deverbal nouns like *drunkard* and *singer* is syntactically derivable and need no morphological characterization. Denominal nouns, de-adjectival adjectives and deverbal verbs are by definition absent in the class of nouns, adjectives and verbs resulting from the transposition of nominal, adjectival and verbal predicates.

None of the constructions within the frame of (3)(c) is discussed by Williams 1981. This is to be expected, because in most of the cases we have to do with the transposition of a predicate from one category into another, so that preser- vation of argument structure is to be expected. There may be deletion of one or more arguments as discussed above in relation to the non category changing suffixations of (3)(a).

A problem rises however in relation to the nouns *drunkard* and *singer* in (6) and their argument structure. Williams (1981) assumes that such nouns have an external argument in all their uses, including referential, non copulative ones. In that case (6) represents uniformly the transposition of nominal adjectival and verbal predicates into predicates of one of the other categories, and examples (6)(c) and (e) must be read with the parenthesized words.

But I do not consider this to be the case. That is why I described *drunkard* and *singer* differently in the foregoing, supposing the examples (6)(c) and (e) to be read without the parenthesized words. This problem concerns the notion

of external argument. And as it is more convenient to discuss this in relation with constructions like *reader* 'book to be read' and *handleable*, I will postpone the discussion until we reach those constructions in section **6**.

Whatever the result of that discussion, we have seen here that the constructions within the frame of (6)(c) extend systematically over the categories N, A and V as input as well as output.

4. Let us discuss now the constructions illustrated in the upper line of (3)(c), indicated there as particular because of the "possessive" role of their base noun. I will mainly discuss the simple constructions of *villager*, *gloomy* and *agonize*. I consider *to hammer* to be a causative counterpart of *to agonize*, and causative constructions in general will be discussed in the next section.

In terms of X-bar structure we find here a nominal base as complement and a nominal, adjectival or verbal suffix as head.

As to predication, Williams (1981) does not consider these constructions. But if we want to find out how the argument structure of these constructions is to be described, it is useful to see first what is going on semantically. It turns out that all these denominal constructions contain the meaning element 'in relation to':

(7) (a)	*villager*	'somebody in relation to village'
	homer	'something in relation to home'
(b)	*gloomy*	'in relation to gloom'
(c)	*agonize*	'to be in relation to agony'
	hammer	'to cause to be in relation to hammer'

As the examples above show, the particular meaning of such a denominal noun, adjective or verb can be specified on the basis of the meaning of the base: 'somebody in relation to village' is 'somebody living in a village', etc.

This semantically very general relation 'in relation to' is exactly what Williams (1981, 88—89) describes — for syntax — as characteristic of a "possessive" argument, which one finds, for example, in *John's book*: "... a possessive NP may bear any relation whatever to the head noun: this is a great exaggeration, but it is a first approximation that is difficult to improve upon". It is the relation described by Allen 1978 for compounds as relation R, What we find now is that the creation of such an argument position is a possible operation on the level of argument structure in derivational morphology.

What exactly is going on in syntactic possessive NPs is a matter of debate. Stowell (1983, 292—296) considers *John's* in *John's book* to be a subject of NP. But this implies a theta role of a particular kind (294): "If lexical NP appears as the subject of such an NP, it is assigned a theta role T involving

some connection with the head. . . . T can mean almost anything at all: John's book could be the book which John owns, the book which John wrote, the book which discusses John etc." Moreover it necessitates the introduction of "an auxiliary principle P that T may not be assigned to PRO". According to Williams (1981) it is not a subject or external argument, nor is it according to Zubizarreta (1985, 257): ". . . the determiner position [*John's* in *John's arrival*] is not the subject of the NP. In fact, the determiner can bear any type of semantic relation to the head of the NP. It can specify the time of the action, the Possessor, the Agent, or the Theme, . . . ".

However this may be in syntax, in morphology the conclusion is inevitable that derivation (as well as compounding) constructions can create such an argument position, in the first place with nominal suffixes like *-er* in *villager* or *homer* (or with nominal lexical heads as in *home-run*). But then there is no reason why adjectival and verbal suffixes would not have the same possibility, as in *gloomy* or in *to agonize* and in causative *to hammer*.

Given the particular argument position and the fact that the role of arguments is performed by nominal expressions only, it is to be expected that these constructions can only imply nominal bases, and so they do. And they do it, as we have seen, systematically with nominal, adjectival and verbal suffixes.

5. Let us finally discuss, in this section and the next one, the righthand column of (3)(c) and the two bottom lines. They have all to do with verbs, either as derived words with a verbal suffix as head or as bases of derived words. These are the cases where, apart from the abstract nouns in (3)(b), externalization and internalization play an essential role. And this is to be expected for these verbal or deverbal constructions, because externalization and internalization have to do with different ways of organizing arguments around verbs mainly.

Let us first consider the righthand column of (3) (a), which contains the causative counterparts of the constructions in the second column from the right:

(8) (a) *agonize* 'to be in relation to' — *hammer* 'to cause to be in relation to'
(b) *head* 'to be' — *mummify* 'to cause to be'
(c) *idle* 'to be' — *harden* 'to cause to be'
(d) (*to run*) — *run* 'to cause to run'

In terms of X-bar structure, these constructions contain a verbal suffix as head and a nominal, adjectival or verbal base as complement.

As for argument structure, we find here a kind of internalization which is different from the one which we find with abstract nouns. In contrast with

those, we find here not only an external argument becoming internal, but at the same time the creation of a new ordinary external argument:

(9) *the material is hard → the temperature hardens the material*

Discussing the notion of external argument in the next section, we will see that in the case of abstract nouns we only find the creation of a particular kind of external argument R in Williams' terms, or — more probably so in my view — of no external argument at all. The kind of argument structure attributed here to the cases illustrated in (8) and (9) is to be expected given the possibility to derive causative verbs.

And we see once more that the constructions discussed here extend systematically over the possible categories: nominal, adjectival and verbal bases allow causative verbal constructions, non derived verbal bases like *run* included. The case of *to hammer* combines, as its position in the matrix of (3)(c) makes us expect, the causative construction with the 'possessive' construction.

6. In a way we see the reverse of this in the two bottom lines of (3) (c). There we find verbal bases for "passive" nouns, adjectives and verbs. These constructions are exemplified in (10):

(10) (a) (*somebody*) *reads the book* → (*the book is a*) *reader*
(*somebody*) *slips into the shoe* → (*the shoe is a*) *slipper*
(b) (*somebody*) *handles the box* → *the box is handleable*
(*somebody*) *relies on John* → *John is reliable*
(c) (*somebody*) *melts the snow* → *the snow melts*

In terms of X-bar structure, all these constructions contain a nominal, adjectival or verbal suffix as head.

As for argument structure, the simplest case is the construction illustrated by *handleable*, described by Williams as presenting externalization of an internal argument, as in (10)(b). *Reliable* shows that it is not always the direct object which is externalized.

According to Keyser—Roeper 1984, ergative verbs like *to melt* are comparable, in the sense that they are derived from the corresponding transitive verbs with externalization of an internal argument, as in (10) (c). Williams (1981, 99) supposes that the inverse is taking place, to wit derivation of transitive *melt* from intransitive *melt* with internalization of the external argument, as we supposed in the preceding section for *run*. A more detailed discussion about zero derivation and conversion will have to decide about these proposals and a third one given in Lieber (1980). According to Lieber we have to do in these cases with conversion as a non derivational lexical relation.

This does not invalidate my reasoning, because overt suffixation for the verbal constructions of the kind discussed here is in principle possible and in fact occurs in other languages. Moreover, argument structure manipulations like the ones we consider here must relate to conversion as well as to affixation.

The case of nouns such as *reader* 'book to be read', *slipper* 'shoe you can slip into', confronts us with Williams' problematic treatment of the notion external argument. According to Williams (1981, 86) these cases present externalization also. Concerning cases like *the book* in *the book is a reader* or *that* in *I consider that destruction of a city by evil forces* 'I consider that to be . . .', he says: "the external argument of . . . a noun has no counterpart in the verbal system; suppose we invent the label *R* to name that argument of the noun which is external. Then we would assign *destruction* the argument structure (*R*, Actor, Theme). The label R is meant to suggest 'referential', since it is the argument position R that is involved in referential uses of NPs as well." In my view, two levels of reasoning are mixed up here, to wit the level of argument structure and the level of reference. In Travis—Williams (1983), for that matter, we find a statement which is a little weaker, restricting the external argument role to the case where it relates to a predicative noun: " . . . NPs do have external arguments, at least when they occur as the head of an NP that is used predicatively, as in *John is a fool.*" .

In relation with *reader* and *singer*, remember also the case, encountered in section **3**, of 'active' deverbal nouns like *singer*, where a comparable problem arises. According to Williams, *singer* also has an external argument, even in its referential, non copulative uses.

In contrast with Williams, Fabb (1984, 206—210) denies the existence of such external arguments R when he says about deverbal nouns like *singer* and *reader*: " . . . the external theta-role is assigned to the affix *-er*." This role can be agent (*gambler*), source (*gusher*), theme (*twister*), benefactor (*hearer*), etc. In giving this analysis, Fabb denies that such nouns have an external argument in all their uses.

I agree with him, but I think that something more can be said on the basis of Chierchia's analysis of predication. I feel that in words like *singer* and *reader* the external argument is, in conformity with Fabb, assigned to the suffix: a *singer* indicates an individual who is the external argument of *sing*. But in their secondary, predicative use such nouns can be linked by means of a copula to an external argument which performs the same theta-role: *John is a singer*. This analysis does not mix up, like Williams', the level of argument structure and the level of reference. Yet it maintains a close relationship between preservation of argument structure or externalization in cases like *fidgety* and *handleable* on the one hand, and morphological incorporation à la Fabb accompanied or not by a simile of preservation or externalization in the cases of *singer* and *reader*.

As to the entire set of 'passive' constructions discussed in this section, we see once more that they extend in a systematic way over the three possible categories N, A and V.

7. Before concluding this article, it may be useful to note following. The constructions discussed here are defined in terms of X-bar structure as well as of argument structure manipulations. But this does not mean that these manipulations are restricted to these constructions. Thus, if the English verb *ennoble* consists of a verbal prefix as head and an adjectival base as complement, as Selkirk (1982, 87—89) suggests, it represents a prefixal construction of causative de-adjectival verbs related to the suffixal causative construction of *harden*, and it has the same argument structure manipulation. And in so far as a language presents conversion in its lexicon, this can be associated, as I suggested in the preceding section, with the same kinds of argument structure manipulation as affixal constructions.

The example of *ennoble* shows also that in the domain of prefixation one can distinguish constructions in the same way as we have done for suffixation. *Ennoble* and *becalm*, for example, belong to the same prefixal construction notwithstanding their different individual suffixes. Moreover, it will be clear that the kind of analysis which I advocate sets forth the kinship of affixation with compounding, whose constructions are not fundamentally different. Thus, the compound *home-run* is comparable to derivatives like *homer* and *villager*, the only difference being that the nominal head in the compound *home-run* is the noun *run* instead of the nominal suffix *-er*.

This means that the conclusions concerning suffixation which we will draw in the next section point to a considerable extent to the larger domain of derivation and even of morphological constructions in general. On the other hand, as I suggested at the end of section **0**, they must to a large extent concern very general, non language-specific regularities of morphological structure.

8. Against this background, the preceding sections allow us then to draw the following conclusions concerning English suffixation. English has a limited number of suffix constructions, which are almost entirely the result of the restrictions of X-bar structure and of argument structure.

Within the restrictions of X-bar structure, the constructions discussed here turn out to extend systematically over the three categories of morphological structure N, A and V.

The restrictions at the level of argument structure can to a large extent be derived from the formal semantic theory of predications as proposed in Chierchia (1985). As for the possible manipulations of argument structure proposed by Williams (1981), we have found out that the notion of external argument has to be modified. Moreover we must add as possible manipulations the creation of a "possessive" argument and morphological incorporation.

References

Allen, Margaret R.: Morphological Investigations. PhD dissertation University of Connecticut, Storrs Ct. 1978.
Carrier-Duncan, Jill: Linking of thematic roles in derivational word formation. In: Linguistic Inquiry 16 (1985), 1—34.
Chierchia, G.: Formal semantics and the grammar of predication. In: Linguistic Inquiry 16 (1985), 417—443.
Fabb, N. A. J.: Syntactic Affixation. PhD dissertation, MIT 1984.
Keyser, S. J. and Roeper, T.: On the middle and ergative constructions in English. In: Linguistic Inquiry 15 (1984), 381—416.
Lieber, Rochelle: On the Organization of the Lexicon. PhD dissertation, MIT 1980.
Selkirk, Elisabeth O.: The syntax of words. MIT Press, Cambridge, Ma. 1982.
Stowell, T.: Subjects across categories. In: Linguistic Inquiry 1 (1983), 81—114.
Travis, L. and Williams, E.: Externalization of arguments in the Malayo-Polynesian languages. In: LR 2 (1983), 27—47.
Williams, E.: Argument structure and morphology. In: LR 1 (1981), 81—114.
Zubizarreta, Maria Luisa: The relation between morphophonology and morphosyntax: The case of Romance causatives. In: Linguistic Inquiry 16 (1985), 247—289.

Acta Linguistica Academiae Scientiarum Hungaricae, Tomus 36 (1–4) pp. 273–276 (1986)

CRITICA

Lou Boves: The Phonetic Basis of Perceptual Ratings of Running Speech.
Floris, Dordrecht 1984, 188 pp.

Phonetics is an interdisciplinary science for its subject, speech, should be examined from several aspects. This commonplace has been well known from the very birth of the discipline but the proper place of phonetics, especially within linguistics, has been open to question for centuries. The commonly used term "linguistic phonetics", by way of ringling out one aspect, points to the presence of phonetic results and methodology in many other scientific and engineering fields.

Our knowledge concerning human speech is uneven. On the one hand, there are a lot of reliable results, especially in speech articulation and in acoustics, and a lot of uncertainty and inconsistency, on the other, in psychoacoustics, speech perception or sociophonetics. The monograph under review makes an attempt to solve this contradiction in its topic by establishing links between the aspects examined separately. The author tries to analyze the running speech which is a rare subject in the existing literature. The book deals with vocal aspects of speech and their effects on "person perception". The declared aim of the author is to examine how naive listeners perceive speech and in what way their perception of somebody's speech influences the perception of the person's personality. The aim of this research was to explain perceptual ratings of vocal speech characteristics by means of acoustic measurements. The opposite research programme, that is obtaining perceptual data for different speech-like stimuli and drawing conclusions concerning the processes of perception or the impact of acoustic properties of the stimuli on identification procedure, is better known. In the present monograph the author undertakes a pioneer examination when he makes an effort to provide correlations between the results of an evaluation experiment and of acoustic/articulatory measurements. The chapters dealing with these correlations are new. The articulatory and acoustic analyses, however, have good traditions but I am not sure that they concentrate on running speech as a whole, rather on larger or repeated parts of running speech (e.g. vowels or voiced intervals).

There are 7 chapters in all, containing the theoretical/methodological/technical background, the perceptual experiments, the articulatory and acoustic measurements and the consequences. The initial chapter is entitled "Research into non-verbal aspects of speech" (1—11). That "non-verbal" communication is taken to be identical with vocal aspects seems to be disputable to the reviewer.

The second chapter, "The construction of a scaling instrument" (11—27), deals with the development of an instrument for obtaining perceptual ratings ("semantic differential" (20)). It consists of 35 bipolar scales, e.g. "colourless

— sonorous", "pleasent — unpleasent", "loud — soft", "creaky — not creaky", "regional — not regional", "expressive — expressionless", "polished — slovenly", "quick — slow", "agitated — calm" etc.

The third chapter, "The evaluation experiment" (29—58), discusses a number of properties of the semantic differential, first of all the linearity and the relevance of scales. The author carries out careful experiments on the evaluation of running speech in terms of the dimensions labelled "Voice dynamics", "Articulation quality", "Pitch level/Voice colour", "Tempo" and a factor stressing the socio- and paralinguistic aspects of the rating form. The stimuli to be judged were recordings of eleven subjects who read a neutral and a non-neutral text. These texts were tested by 352 listeners (undergraduate students from Amsterdam and Nijmegen). In order to obtain significant results statistical criteria were used.

The next chapter is entitled "Separation of the contributions of phonation and articulation to the radiated speech signal" (61—107). Findings presented here confirm the author's belief that phonation and articulation should be distinguished and can, in fact, be separated. Data were obtained from electroglottographic and photoglottographic measurements, and from the procedure relating to subglottal pressure. Valuable and important results exist about this problem in phoniatric literature.

The fifth chapter is devoted to the development of acoustic measurements planned with the purpose of explaining perceptual ratings (107—135). A large number of acoustic measurements have been described by means of essentially automatic processing techniques. Some of them are generally used in acoustic investigations; whereas others are less known and provide interesting and new results with regard to running speech. However, I share the opinion of the author when he writes: "This is an area where more research is necessary to answer all questions" (124).

An attempt is made to explain "Relations between acoutic measures and perceptual ratings" in the sixth chapter (135—165). There is no doubt that the correlation of the "subjective" judgements (in spite of the accurate mathematical-statistical background) and the 'objective' data is not a simple operation. The results, however, seem to be useful and reliable due to the correct manipulations of the data. On the basis of the estimations a number of hypotheses and methodological refinements have been proposed to arrive at a better understanding of the remaining questions or factors. (The factor "speakers" proved to be significant at the 1% level for all the pitch measurements, whereas the factor "text" never reached the 5% significance level.)

The last chapter is entitled "The status quo and perspectives" (165—173). The author's over-modest conclusion is that his work tries to bring the problems of the investigated domains a little closer to a solution. Although he has no explanations for some observations and results obtained (e.g. correlation of formant analyses and the ratings by his semantic device), his book is successful as a general account of an almost unknown and little investigated field in phonetics.

Finally, let me make a few remarks concerning details of lesser importance. It is not adequate to use read texts for testing. It would have been closer to normal communication to make recordings from spontaneous speech. The high-level technical and engineering basis seems to be opposed to the tempo-measurements which were performed by handclocking of a stopwatch. It is likely that the legends of figures 4.11 and 4.12 have been mixed up (pp. 84 and 89). The mathematical explanations and the description of some techniques require too high a knowledge, some parts are not easy to follow for the traditionally qualified speech scientists.

The book contains valuable material; I think the results are a great step forward in the interrelation of different levels of

investigation in speech science. Besides, the results can be applied to other disciplines too, as, e.g. in phoniatrics, logopedics, sociolinguistics; they can also be used in the research dealing with automatic speaker recognition.

Mária Gósy

Dictionary of American Regional English, Volume I, Introduction and A—C. Chief Editor: Frederic G. Cassidy

The Belknap Press of Harvard University Press, Cambridge, Mass. 1985, clvi + 903 pp.

Although the history of the Dictionary of American Regional English (DARE) dates back to 1889, the founding year of the American Dialect Society, it was only in the early 1960s that the project gathered considerable momentum. Over 20 years of hard work by Professor Cassidy and a host of other scholars has now borne fruit: the first of the five volumes planned. DARE was much awaited and its appearance was duly celebrated. When Volume I appeared the governor of Wisconsin proclaimed November 16, 1985 DARE Day to honor Cassidy and his fellow scholars at the DARE center in Madison, Wisconsin. The jacket informs us that in the New York Times William Safire called DARE "the most exciting linguistic project going on in the United States", while the New Yorker called it "an alphabetized, highly detailed history of the United States".

The introductory material comprises the following parts: Introduction (xi—xxii), The DARE map and Regional Labels (xxiii—xxxv), Language Changes Especially Common in American Folk Speech (xxxvi—xl), Guide to Pronunciation (xli—lxi), Text of Questionnaire (lxii—lxxxv), List of Informants (lxxxvi—cli) and a List of Abbreviations (clii—clvi).

A great deal of the material that DARE contains was gathered by over eighty fieldworkers in 1,002 communities throughout the United States between 1965 and 1970. The questionnaire contained 1,847 questions. Five community types were established: urban, large city, small city, village and rural. "The choice of informants was generally balanced with an eye to [. . .] community type, sex, race, age, and education — but with a deliberate weighting toward older people" (xiv). The approximately 2,500,000 answers which the 1,002 completed questionnaires yielded form the core of the computerized DARE corpus. In addition, a large body of oral and written, published and unpublished materials has been used from Dialect Notes through the aluminum disks of New England speech recorded by Miles L. Hanley and Guy S. Lowman, Jr. in 1932—1934, to letters from the public and folk names of plants and animals.

A precious part of the DARE collections are the 1,843 audiotapes recorded during the fieldwork. Each recording contains about 20 minutes of free speech on a familiar topic plus a reading of the children's tale "Arthur the Rat" in DARE's revised version. Thus, the audiotapes "constitute a unique record of American pronunciation, drawn from many levels of life and all fifty states" (xiv). The tapes were studied for regional pronunciation and lexical features which might be entered into the DARE files.

For the purposes of DARE the term r e g i o n a l is defined on p. xvi as (1) any word or phrase whose form or meaning is not used generally throughout the United States but only in part (or parts) of it, or by a particular social group, and (2) any word or phrase whose form or meaning is distinctively a folk usage (regardless of region). F o l k u s a g e is characterized as "that which is learned in the home and in the community, from relatives and friends, not from schooling, books, or other outside forms of communication" (ibid.).

Each DARE entry has three parts. The first conveys the basic information: headword or -words, part of speech,

pronunciation, variant forms, etymology, geographic range, usage, cross-references, and editorial notes. The second part contains meanings which (if there are more than one) are numbered although standard senses are not normally treated. The third part of an entry contains dated quotations to support the definitions. "The editors — the Introduction (p. xx) says — have sought to give the earliest example found in American use, at least one per century, and a late or recent one."

A highly original feature of DARE is its use of computer-produced maps. A DARE map illustrates the regional distribution of words and phrases elicited by the questionnaire. Though it resembles the conventional areal map of the USA, it appears distorted because it displays population density (or the number of DARE informants, but that is roughly proportional to the state populations) rather than land area. This makes Nevada, for instance, a tiny state on the map but New York a large one. With the help of the DARE maps and the descriptive labels, a unique feature in linguistic geography — a view of the *nationwide* distribution of American speech — emerges. Thus the entry *beau dollar* 'a silver dollar', for instance, contains the usage label *esp freq among Black speakers*, which was arrived at by analyzing the responses in terms of the race of the informants: 27 of the 34 informants who used the expression were Black although Blacks make up a much smaller percentage (6.7%) of the entire sample of DARE informants. The map shows that the expression is used chiefly in the South and the South Midland. When informants and their race are plotted on the map, "it reveals that the spread of this term northward is the result of the migration of southern Blacks into the urban North" (xxx).

It would be impossible for the present reviewer to do justice to DARE even if he had ten times as much space as he has been allotted. DARE is a truly monumental work in American English and universal dialectology and lexicography. It displays thorough scholarship, contains an incredibly rich store of information, and breaks new ground in dictionary-making. Study it for yourself.

M. Kontra

Are you studying Hungarian? If so, you're likely to want to know the inflectional possibilities of the 70,000 headwords in the *Magyar Értelmező Kéziszótár* (Hungarian Explanatory Dictionary). So you need the companion volume

Magyar ragozási szótár — *Dictionary of Hungarian Inflections*

The Hungarian and English introductions in this dictionary provide all the necessary information about the 12 basic paradigms (6 conjugations and 6 declinations) and the non-basic paradigms, in tables which indicate the respective vowel harmony classes.

Hungarian is the most agglutinating of the European languages, and it has a very elaborate system of inflection and derivation. Of the several hundred suffixes forms of a headword, there are today 66 verb-forms and 33 noun-forms which are characteristic of the root words.

The Dictionary of Hungarian Inflections gives the declinations of adjectives, the possible degree forms of adjectives, and also indicates words with defective paradigms. Boundaries within compound words are indicated and a special sign indicates the separable verbal prefixes.

Planned publication date: December 1993.

Orders should be placed with the Linguistics Institute, Hungarian Academy of Sciences, H-1250 Budapest, P.O. Box 19., Hungary.

PRINTED IN HUNGARY

Akadémiai Kiadó és Nyomda Vállalat, Budapest

the publisher, the place and date of publication as well as the number of pages), the author's full name and address and an abbreviated title of the paper (not exceeding 45 characters including spaces). The contribution proper should begin on page 2.

Tables, diagrams and illustrations (with the author's name and an Arabic number) should be presented on separate sheets. Captions should be typed on a separate sheet and placed at the end of the manuscript.

Footnotes should be typed double-spaced on a separate sheet at the end of the article.

Please underline all emphasized words with a broken line (— — — —). Use double underlining for bold face (e.g. titles of sections). All language examples (and only these) are to be italicized (single underlining). Do not use capitals, small capitals or bold face for emphasis, authors' names or quotations.

Citations in the text should be enclosed in double quotation marks (" " in case of a paper written in English, „ " in German and « » in French or Russian).

Reference to a publication should be made by the name of the author, the year of publication and, when necessary, page numbers in the following ways:

... as described by Schmidt (1967) ...
... as referred to by Hoover (1967, 56–78; 1976, 43).
... mentioned by several authors (Elgar 1978, 77; Williams 1981, 154–156) ...

An alphabetically arranged list of references should be presented at the end of the article, following the footnotes. If there are two or more books or papers published by the same author in the same year, lower case letters a, b, c etc. should be added to the year of publication both in the list of references and in the text proper. Avoid using abbreviations in the references; if they are essential, a list of explanations should be provided. For books, please include the name of the publisher, the place of publication and the year; for articles in journals, please include volume, year and page numbers in the following fashion:

Bárczi G.: Magyar hangtörténet² [The history of Hungarian sounds]. Akadémiai Kiadó, Budapest 1958a.
Bárczi G.: A szótövek [Word stems]. Akadémiai Kiadó, Budapest 1958b.
Lakoff, G.—Peters, P. S.: Phrasal conjunction and symmetric predicates. In: Reibel, D.–Schane, S. (eds): Modern Studies in English. Prentice-Hall, Englewood Cliffs 1969, 113–141.
Ross, J. R.: Constraints on Variables in Syntax. Ph. D. dissertation, MIT 1967.
Sauvageot, A.: Le problème de la parenté ougrienne. In: Etudes Finno-ougriennes 13 [1967], 123–142.
Vasmer, N.: Etimologičeskij slovar' russkogo jazyka IV. Nauka, Moscow 1973.

For marking subsections decimal notation should be applied. Do not use more than four digits if possible.

Titles of sections and their decimal notations should be marked by double underlining. Between sections leave out double the usual space.

Examples within the text of the article should be marked by single underlining, otherwise they are to be left unmarked and separated from the body of the text by spacing and placing them in a new paragraph. Meanings are to be rendered between inverted commas (' '). If glosses are given morpheme by morpheme, the initial letter of the gloss should be placed exactly below that of the example. Grammatical morphemes can be abbreviated in small case letters connected to the stem or the other morphemes by a hyphen. No period should be applied in such abbrevations. For example:

(1) (a)	A	sólymaid	elszálltak
	the	falcon-gen-pl-2sg	away-flew-3pl
	'Your falcons have flown away'.		

Examples can be referred to in the text as (1a), (1a—d), etc.

Authors are requested to check their manuscripts very carefully as corrections in proof stage cause delay and are expensive. One proof will be sent to the author. Please read it carefully and return it by air mail within one week to the editor: Acta Linguistica, MTA Nyelvtudományi Intézet, H-1250 Budapest, P. O. Box 19.

HU ISSN 0001-5946

Index: 26.018